Universidade de Aveiro Departamento de Matemática
2012

Duarte
Polónia Rodrigues

Applications of Systems Theory in the Theory of Codes

Aplicações da Teoria dos Sistemas á Teoria dos Códigos

Duarte
Polónia Rodrigues

Applications of Systems Theory in the Theory of Codes

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Matemática e Aplicações, realizada sob a orientação científica de Paolo Vettori, Professor Auxiliar do Departamento de Matemática da Universidade de Aveiro

o júri / the jury

presidente / president	**Doutora Maria Raquel Rocha Pinto** Professora Auxiliar da Universidade de Aveiro
arguente / examiner	**Doutor Adão Paulo Soares Silva** Professor Auxiliar da Universidade de Aveiro
orientador / supervisor	**Doutor Paolo Vettori** Professor Auxiliar da Universidade de Aveiro

agradecimentos

Um agradecimento especial ao meu orientador, pela forma especial como me conduziu neste processo. Obrigado pela dedicação, pela ajuda e por tudo o que me ensinou.

Obrigado também ao professor Adão Silva, por mais uma vez atender ao meu pedido e dedicar parte do seu tempo a ler esta tese e estar presente na sua apresentação.

Um agradecimento também á minha familia que nunca me faltou com o seu apoio, e aos elementos da secretaria do departamento de matemática, em especial á Sandra, por me ajudarem e aturarem durante todo o processo.

palavras-chave

Álgebra, Teoria dos Códigos, Códigos de Reed–Solomon,
Teoria dos Sistemas, Abordagem Comportamental,
Interpolação Polinomial Bidimensional

Resumo

Os códigos the Reed Solomon são códigos cíclicos não binários com símbolos de código num corpo de Galois. Eles foram descobertos em 1960 por L. Reed e G. Solomon. Nas décadas após a sua descoberta, os códigos de RS gozaram de inúmeras aplicações. Este trabalho é maioritariamente focado nos algoritmos de codificação e descodificação dos códigos RS, analisando uma relação entre a teoria dos códigos e uma técnica de identificação desenvolvida no âmbito da abordagem comportamental á teoria de sistemas.

keywords Algebra, Coding Theory, Reed–Solomon Codes,
Systems Theory, Behavioral Approach
Bidimensional Polinomial Interpolation

Abstract The Reed–Solomon codes (RS codes) are non-binary cyclic codes with code symbols from a Galois field. They were discovered in 1960 by I. Reed and G. Solomon. In the decades since their discovery, RS codes have enjoyed countless applications. This work is mainly focused on encoding and decoding RS codes algorithms, analyzing a relation between coding theory and an identification technique developed within the behavioral approach to systems theory.

Contents

Introduction

Coding theory is the study of the properties of codes and their fitness for a specific application. It emerged following the publication of Claude Shannon's of 1948[24]. However, his work was about channel properties and he did not tell how to find suitable codes. There are two types of codes: **Data Compression** (or, source coding), and **Error-Correction** (or, channel coding). When data compression is used in a data transmission application, the goal is speed. The objective is to minimize the amount of data to be transmitted in order to increase the data speed of the transmission. In the error-correction, such the name suggests, the goal is to verify data transmissions by locating and correcting transmission errors. This thesis will focus on the second type of codes. There are two classes of error-correction codes: the **linear** codes and the **non-linear** codes.

Both of these classes of codes allow for efficient encoding and decoding algorithms, however very little is known about the properties of non-linear codes, as we can read in[4]. In this work we will use only linear codes. Linear codes are partitioned into **convolutional codes** and **block codes**. The first block code was introduced in 1950 by Hamming[2], and it is called **single-error-correcting** block code. Later in 1954, Muller invented the class of **multiple-error-correcting** codes and Reed gave a decoding algorithm for them[21]. The major advances came with Bose Ray-Chaudhuri (1960)[3] and Hocquenghem (1959)[5], when they found a large class of **multiple-error-correcting** codes, the **BCH** codes, and with Reed and Solomon (1960) and Arimoto (1961)[22], when they discovered another class of multiple-error-correcting codes called **Reed-Solomon** codes RS. Our main work will be concerned with RS codes.

The discovery of BCH codes lead to a search for practical methods of designing the hardware and software to implement the encoder and decoder of these codes. The first good algorithm was found by Peterson (1960)[19]. Later, a powerful algorithm for decoding was discovered by Berlekamp and Massey (1960)[9]. In a more recent work[27], Sudan introduced a RS

decoding method that outperforms all existing ones.

In 1967, Massey and Sain published[18], where they start to establish one relation between the area of coding theory and the area of linear systems theory. In 1995, York and Rosenthal presented this relation as a particular case of the **behavioral approach**[23]. It is shown in [9] how the theory on behavioral modeling leads to a transparent interpretation of various existing decoding methods as well as to the derivation of an insightful decoding algorithm.

In particular, the Berlekamp–Massey algotithm is interpreted as behavioral modeling for single-input-single output partial realization, as it was presented in [16]. Also a multivariate version of this algorithm was builded in[6], and applied in[7]. In this work we will be concerned in the above mentioned result of Sudan [27]. We summarize the idea of[13], where Sudan's approach is interpreted as a behavioral modeling for multi-variable interpolation.

It is also important to refer one common difference between the coding theory and the system theory which is the **alphabet** used in each one. In code theory is usual to use finite alphabets whereas in system theory the more usual to use infinite. We will show the implications of finite fields for behavior modeling.

The structure of this thesis is the following:

Algebraic structures In this section we present the basic algebraic concepts which will be used along this work, such as equivalence classes, cyclic groups and quotient rings.

Finite Fields In this section we further investigate finite fields and their structure, in particular the concept of extension fields and their construction. It has special importance because all the "numbers´´ in code theory are elements in finite fields.

The behavioral Approach to Systems Theory This section starts with the definition of a dynamical system and its properties and we define autoregressive systems. Then we describe the concept of mathematical model, as it was introduced by Jan. C. Willems in [31], and it important application in a dynamical systems theory. We present the algorithm to find the Most Powerfull Unfalsified Model (MPUM) for a dynamical system.

Codes In this section we describe the general characteristics of the linear block codes, as the concept of distance and error correcting capability, and we introduce the cyclic codes. In

particular, we describe the Reed-Salomon codes. This is one strong class of the linear cyclic codes, whose construction is made over elements of the finite fields introduced in section 2, and we will apply this codes in a practical situation in section 5.

Decoding BCH and RS Codes In this last chapter are presented some algorithms to decode RS codes, giving a special fucus to the behavioral decoder. In this subsection we recall to the idea of Sudan, which consists in the construction of an interpolating polynomial [27]: this task will be accomplished by using techniques developed in the framework of the behavioral approach, as it is presented by M. Kuijper in [15]. The idea is to associate a set of trajectories to the received data and then apply the behavioral theory presented in[30] by Willems, and find the MPUM for this trajectories, which is the smallest polynomial matrix in the shift whose kernel represents the behavior generated by those trajectories. With the row of minimal weighted degree of this matrix we construct an polynomial which interpolates the original data points.

Chapter 1

Algebraic Structures

In this chapter we introduce some basic algebraic concepts and theorems, which will be used in this thesis.

1.1 Functions and equivalence relations

Definition 1.1: Given two nonempty sets S and T, any subset of their cartesian product $R \subseteq S \times T$ is a (binary) **relation**. If $(x, y) \in R$, we will write also xRy.

A function is a special type of relation.

Definition 1.2: A relation $R \subseteq S \times T$ defines a **function** if it satisfies the following conditions:

- for every $x \in S$ there exists $y \in T$ such that $(x, y) \in R$;
- if $(x, y_1) \in R$ and $(x, y_2) \in R$, then $y_1 = y_2$.

A special notation is used for functions: the sets S, T, and R are called, respectively, **domain**, **codomain**, and **graph** of the function. If the name of the function is f, we will write $f : S \to T$ and $f : x \mapsto y$ or $y = f(x)$ whenever $(x, y) \in R$.

Observe that the notation $y = f(x)$ makes sense, since every x in the domain is associated with a unique y in the codomain, which is called **image** of x.

The following properties of functions will be used frequently.

Definition 1.3: A function $f : S \to T$ is

- a **surjection** or **surjective** if

$$\forall y \in T, \exists x \in S : f(x) = y;$$

- an **injection** or **injective** if

$$\forall x_1, x_2 \in S : f(x_1) = f(x_2) \Rightarrow x_1 = x_2;$$

- a **bijection** or **bijective** if it is both a **surjection** and an **injection**.

This chapter is devoted to the definition and characterization of different algebraic structures that a set may exhibit. When a function is 'compatible' with the domain's and codomain's structure some special names are used.

Definition 1.4: An **homomorphism** is a structure-preserving function between two sets equipped with some algebraic structure. A homomorphism is called

- **monomorphism** when it is injective,
- **epimorphism** when it is surjective, and
- **isomorphism** when it is bijective.

Two sets A and B are **isomorphic**, denoted by $A \cong B$, when there exists an isomorphism between them.

More specific properties of homomorphisms will be given for each algebraic structure that will be introduced starting from next section.

Another very important type of relation is introduced in the following definition and will be widely used throughout this thesis.

Definition 1.5: An **equivalence relation** is a binary relation on a set S, i.e., $\sim \subseteq S \times S$ satisfying three properties:

- (reflexivity) $\forall a \in S, a \sim a$
- (symmetry) $\forall a, b \in S$, if $a \sim b$ then $b \sim a$
- (transitivity) $\forall a, b, c \in S$, if $a \sim b$ and $b \sim c$, then $a \sim c$

Definition 1.6: Given a set S and an equivalence relation $\sim$ on S, the **equivalence class** of an element a in S is the subset

$$\overline{a} = \{x \in X : x \sim a\} \subseteq S.$$

The set of all equivalence classes of S will be denoted by $S/\sim$.

Remark 1.7: An equivalence class on S induces a **partition** of S, which is a family of disjoint subsets (the equivalence classes) whose union is S.

Example 1.8: We will prove that the relation $\sim$ on $Q = \mathbb{Z} \times \mathbb{Z} \backslash \{0\}$ defined by

$$(a, b) \sim (c, d) \Longleftrightarrow ad = bc$$

is an equivalence relation. Actually, it is

- reflexive, since $(a, b) \sim (a, b) \Leftrightarrow ab = ba$;
- symmetric, because $(a, b) \sim (c, d) \Rightarrow (c, d) \sim (a, b)$ as long as

$$ad = bc \Leftrightarrow cb = da;$$

- transitive, i.e.,

$$(a, b) \sim (c, d) \text{ and } (c, d) \sim (e, f) \Rightarrow (a, b) \sim (e, f). \tag{1.1}$$

To prove (1.1), note first that, for every $b, d \in \mathbb{Z} \setminus \{0\}$,

$$(a, b) \sim (0, d) \Leftrightarrow a = 0. \tag{1.2}$$

Indeed, since $ad = b0 = 0$ and $d \neq 0$, it follows that $a = 0$; the converse is obvious. Therefore, if $c = 0$ in (1.1), then $a = 0$ and also $e = 0$, being $(a,b) = (0,b) \sim (0,f) = (e,f)$.

When $c \neq 0$, the hypothesis of (1.1) corresponds to $ad = bc$ and $cf = de$. By multiplying left and right members of these equalities, we get

$$adcf = bcde \Leftrightarrow dc(af - be) = 0 \Leftrightarrow af = be \Leftrightarrow (a,b) \sim (e,f),$$

where $d \neq 0$ by definition of Q, thus justifying the second equivalence.

Definition 1.9: For any $a, b \in \mathbb{Z}$ and positive $n \in \mathbb{N}$, a is **congruent** to b modulo n, if the difference $a - b$ is a multiple of n, i.e., $a = b + kn$ for some $k \in \mathbb{Z}$. Equivalently, n **divides** (or is a **divisor** of) $a - b$, denoted by $n|a - b$.

Observe that, once n is fixed, congruence modulo n is an equivalence relation whose classes are

$$\overline{b} = \{a \in \mathbb{Z} : a = b + kn,\ k \in \mathbb{Z}\}.$$

The set of all the equivalence classes is denoted by

$$\mathbb{Z}_n = \{\overline{0}, \overline{1}, \ldots, \overline{n-1}\}. \tag{1.3}$$

1.2 Groups

Definition 1.10: A set G equipped with a binary operation $\diamond$, denoted by $(G, \diamond)$, is a group if the following properties are satisfied:

- (Closure) $a \diamond b \in G, \forall a, b \in G$.
- (Associativity) $a \diamond (b \diamond c) = (a \diamond b) \diamond c, \forall a, b, c \in G$.
- (Existence of Identity) There exists $e \in G$, called **identity**, such that $a \diamond e = e \diamond a = a$, for every $a \in G$.

- (Existence of Inverse) For any $a \in G$ there exist $b \in G$, called **inverse** of a, such that $a \diamond b = b \diamond a = e$.

If the commutative property holds, i.e., $a \diamond b = b \diamond a$, for every $a, b \in G$, the group is **commutative** or **abelian**.

G is a **finite group** if it contains a finite number of elements. The **order** of G, denoted by $\operatorname{ord} G$ is the number of its elements.

We may speak of *the* identity element and *the* inverse of any element in a group, since they are unique, as the following theorems state.

Theorem 1.11: In every group, the identity is unique.

Proof: Suppose that there exist two identity elements e_1 and e_2. Therefore, $e_1 \diamond e_2 = e_2$ and $e_1 \diamond e_2 = e_1$. Thus $e_1 = e_2$. ■

Theorem 1.12: The inverse of each group element is unique, and the inverse of the inverse of a is a.

Proof: Suppose that b and c are two inverse elements of a. Then,

$$b = b \diamond e = b \diamond a \diamond c = e \diamond c = c.$$

Now suppose that d is the inverse of b. So,

$$d = d \diamond e = d \diamond (b \diamond a) = (d \diamond b) \diamond a = e \diamond a = a,$$

and the theorem is proved. ■

Example 1.13: We will prove that $(Q/\sim, \oplus)$ is an abelian group, where the set $Q/\sim$ was defined in example 1.8 and the operation $\oplus$ is defined by $\overline{(a,b)} \oplus \overline{(c,d)} = \overline{(ad + bc, bd)}$.

First of all, observe that $\oplus$ is well defined, since it does not depend on the representatives of each equivalence class. Indeed, if $(a', b') \sim (a, b)$ and $(c', d') \sim (c, d)$, then also

$\overline{(a'd'+b'c',b'd')}=\overline{(ad+bc,bd)}$: by direct calculation,

$$a'd'bd+b'c'bd=b'd'ad+b'd'bc$$
$$d'd(a'b-b'a)=b'b(d'c-cd'),$$

which is true, because $ab'=a'b$ and $cd'=c'd$. As for the group axioms,

- the operation is closed $\overline{(a,b)}\oplus\overline{(c,d)}=\overline{(ad+bc,bd)}\in Q/\sim$, since $bd\neq 0$.
- The identity exists and is equal to $\overline{(0,b)}$, which we will denote by $\overline{0}$, since it does not depend on $b\neq 0$. Actually,

$$\overline{(a,b)}\oplus\overline{0}=\overline{(a,b)}\oplus\overline{(0,d)}=\overline{(ad+b0,bd)}=\overline{(ad,bd)}=\overline{(a,b)},\ \forall d\neq 0.$$

- The inverse of $\overline{(a,b)}$ is $\overline{(-a,b)}$, being

$$\overline{(a,b)}\oplus\overline{(-a,b)}=\overline{(ab+b(-a),bd)}=\overline{(0,bd)}=\overline{0}.$$

Commutativity of $\oplus$ is a direct consequence of the commutativity of the sum of integers.

We will often extend the use of operations to sets and give here a general definition for future references.

Definition 1.14: Suppose that the operation $a\diamond b$ is defined for every $a\in A$ and $b\in B$. Then, we define

$$a\diamond B=\{a\diamond b : b\in B\},\quad A\diamond b=\{a\diamond b : a\in A\}\quad\text{and}\quad A\diamond B=\{a\diamond b : a\in A, b\in B\}.$$

Remark 1.15: When no ambiguity arises, and the operation is clear from the context, we may write just G instead of $(G,\diamond)$. In particular, two types of notation will be used:

Additive group — denoted by $(G,+)$.

The operation '+' is called sum, the identity is $e=0$ and the inverse of $a\in G$ is $-a$. The n-fold composite of $a\in G$ with itself, with $n\in\mathbb{N}$, is called n-th multiple of a and denoted by

$$n\times a=\overbrace{a+a+\cdots+a}^{n}.$$

This operation is extended to $n \in \mathbb{Z}$, being $0 \times a = 0$ and $n \times a = (-n) \times (-a)$ for $n < 0$.

Multiplicative group — denoted by $(G, \cdot)$.

The operation '$\cdot$' is called product, the identity is $e = 1$ and the inverse of $a \in G$ is a^{-1}. The n-fold composite $a \in G$ with itself, with $n \in \mathbb{N}$, is called n-th power of a and denoted by

$$a^n = \overbrace{a \cdot a \cdots a}^{n}.$$

The operation is extended to $n \in \mathbb{Z}$, being $a^0 = 1$ and $a^n = (a^{-1})^{-n}$ for $n < 0$.

Remark 1.16: Note that for any $a \in G$, both the 0-th multiple $0 \times a = 0 \in G$, in the additive case, and the 0-th power $a^0 = 1 \in G$, in the multiplicative case, are equal to the identity element of the group.

In particular, observe that the result of the previous operations is always an element of the group. Especially in the additive case, there is an (intentional) slight abuse of notation, since the same symbol (0) is used with two different meanings ($0 \in \mathbb{Z}$ and $0 \in G$).

Definition 1.17: Consider two groups $(G, \diamond)$ and $(H, \circ)$. Then the function $f : G \to H$ is a group homomorphism if

$$f(u \diamond v) = f(u) \circ f(v), \ \forall u, v \in G.$$

Theorem 1.18: Consider the group homomorphism f of $(G, \diamond)$, with identity e_G, to $(H, \circ)$, with identity e_H. Let u^{-1} denote the inverse in both groups. Then,

1. $f(e_G) = e_H$ and $f\left(u^{-1}\right) = \left(f(u)\right)^{-1}$, $\forall u \in G$ and
2. f is injective if and only if $f(u) = e_H \Rightarrow u = e_G$.

Proof:

1. The first result is a consequence of $f(u)=f(u\diamond e_G)=f(u)\circ f(e_G)$ and of the unicity of the identity. Similarly, the second result follows by the unicity of the inverse, being

$$f(u)\circ\left(f(u)^{-1}\right)=e_H=f(e_G)=f\left(u\diamond u^{-1}\right)=f(u)\circ f\left(u^{-1}\right).$$

2. Applying the first part of this theorem, the 'only if' part is obvious. Vice-versa, to prove the 'if' part, note that

$$f(u)=f(v)\Leftrightarrow f(u)\circ\left(f(v)\right)^{-1}=f(u)\circ f\left(v^{-1}\right)=f\left(u\diamond v^{-1}\right)=e_H.$$

   By hypothesis, last condition implies that $u\diamond v^{-1}=e_G\Leftrightarrow u=v$ and, therefore, f is injective.

■

Definition 1.19: Let $(G,\diamond)$ be a group. If $H\subseteq G$ and $(H,\diamond)$ is a group too, it is called **subgroup** of G.

Theorem 1.20: Let $(G,\diamond)$ be a group. A nonempty subset $H\subseteq G$ is a subgroup of G if and only if $x\diamond y^{-1}\in H$ for every $x,y\in H$, where y^{-1} is the inverse of y (in G).

Proof: '$\Rightarrow$' Since H is a group, for every $x,y\in H$ we have that $y^{-1}\in H$ and thus $x\diamond y^{-1}\in H$. '$\Leftarrow$' Let $y\in H$. If we consider $x=y$, then $x\diamond y^{-1}=y\diamond y^{-1}=e\in H$, so H contains the identity. Therefore, taking $x=e$, $x\diamond y^{-1}=e\diamond y^{-1}=y^{-1}\in H$ and H contains the inverse of its elements. To end the proof, we have to show that the operation is closed in H: for any $x,z\in H$, we showed that $y=z^{-1}\in H$. By Theorem 1.12, $y^{-1}=z$, thus $x\diamond y^{-1}=x\diamond z\in H$. ■

Remark 1.21: Note that, when $H\subseteq G$, $(H,\diamond)$ is a subgroup of $(G,\diamond)$, if and only the identity map $H\rightarrow G$, $u\mapsto u$, is an injective group homomorphism. Often, we will use the latter condition, i.e., the existence of a monomorphism $H\rightarrow G$, to say that H is a subgroup of G even when H is not a subset of G, meaning that H is isomorphic to some subgroup $\widetilde{H}$ of G (which is a subset too).

Example 1.22: The group $(\mathbb{Z},+)$ is a subgroup of $(Q/\sim,\oplus)$, defined in Example 1.13.

Indeed, consider the set $Z=\left\{\overline{(a,1)}\in Q/\sim\right\}$. By Theorem 1.20, it is a subgroup of $Q/\sim$ since, if we consider $\overline{(a,1)}$ and $\overline{(b,1)}$, both in Z, and $\overline{(-b,1)}\in Q/\sim$ is the inverse of $\overline{(b,1)}$, then we have that

$$\overline{(a,1)}+\overline{(-b,1)}=\overline{(a-b,1)}\in Z.$$

Further, the function $\mathbb{Z}\rightarrow Z$, $a\mapsto\overline{(a,1)}$ is a 'natural' isomorphism between $(\mathbb{Z},+)$ and $(Z,\oplus)$. Indeed, it is a homomorphism, since, by definition of $\oplus$,

$$\forall a,b\in\mathbb{Z},\ a+b\mapsto\overline{(a+b,1)}=\overline{(a,1)}\oplus\overline{(b,1)};$$

it is clearly surjective and, by Theorem 1.18, it is injective because $a\mapsto\overline{(a,1)}=\overline{0}\Leftrightarrow a=0$, as we saw in (1.2).

1.3 Cyclic Groups

Definition 1.23: The multiplicative (additive) group G is **cyclic** if there exists $a\in G$ such that $G=\{a^n,n\in\mathbb{Z}\}$, $(G=\{n\times a,n\in\mathbb{Z}\})$. The element a is called **generator**. In general, every $a\in G$ generates a **cyclic subgroup** of G.

The additive group of integers $(\mathbb{Z},+)$ is cyclic, being generated by 1 or by -1.

Example 1.24: The additive group $(Q/\sim,\oplus)$ of Example 1.13 is not cyclic. Suppose that $Q/\sim$ is cyclic and $\overline{(c,d)}$ is its generator. This means that

$$\forall\overline{(a,b)}\in Q/\sim,\exists n\in\mathbb{Z}:\overline{(a,b)}=n\times\overline{(c,d)}.$$

By definition of $\oplus$, it is easy to prove that $n\times\overline{(c,d)}=\overline{(nc,d)}$. Thus, in particular

$$\exists n\in\mathbb{Z}:\overline{(c,2d)}=n\times\overline{(c,d)}=\overline{(nc,d)}\Leftrightarrow cd=2dnc.$$

Once $d\neq 0$, the former condition is equivalent to $c=2nc\Leftrightarrow c(1-2n)=0$.

The solution $c = 0$ is not be acceptable, because the generator would be $\overline{0}$, which is impossible, and the other solution $n = \frac{1}{2}$ is not an integer. So, the statement is proved by contradiction.

Also finite groups may be cyclic or not, as we show in the following examples.

Example 1.25: According to definition (1.3), it is easy to see that $\mathbb{Z}_n$ is a finite group order $\operatorname{ord}\mathbb{Z}_n = n$ with respect to the operation $\oplus$ defined by $\overline{a} \oplus \overline{b} = \overline{a+b}$. Moreover, it is cyclic with generator $\overline{1}$, being $\overline{a} = a \times \overline{1}$ for any $a = 0, \ldots, n-1$.

By finiteness, all the results of the operation can be organized in a table, called **Cayley table**. As an example, the complete structure of the group $(\mathbb{Z}_4, \oplus)$ is

$\oplus$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{0}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{1}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{0}$
$\overline{2}$	$\overline{2}$	$\overline{3}$	$\overline{0}$	$\overline{1}$
$\overline{3}$	$\overline{3}$	$\overline{0}$	$\overline{1}$	$\overline{2}$

Example 1.26: Consider the finite multiplicative group $G = \{1, a, b, ab\}$, such that $a^{-1} = a$ and $b^{-1} = b$. Its Cayley table is the following:

$\odot$	1	a	b	ab
1	1	a	b	ab
a	a	1	ab	b
b	b	ab	1	a
ab	ab	b	a	1

We deduce that this group is not cyclic, since every element $x \in G$, $x \neq 1$, generates (only) the proper subgroup $\{1, x\} \subset G$.

1.4 Rings and Fields

A ring is an abelian group with an additional structure. For instance, the set of integers $\mathbb{Z}$ is a group with respect to addition, but also multiplication can be defined.

Definition 1.27: A set R with two binary operations $+$ and $\cdot$, sum and product, denoted by $(R,+,\cdot)$ is a **ring** if the following properties are satisfied:

- $(R,+)$ is an abelian group.
- Product is closed and associative.
- The distributive rule holds: for all $a,b,c \in R$, $a\cdot(b+c) = a\cdot b + a\cdot c$ and $(b+c)\cdot a = b\cdot a + c\cdot a$.

If also multiplication is commutative, R is a **commutative ring**. If it has a finite number of elements, R is a **finite ring**.

Remark 1.28: Note that in a ring, multiplication does not have necessarily the identity element. However, when it exists, it will be denoted by 1, according to Remark 1.15.

Example 1.29: A non trivial example of abelian ring is the set $P = \{2n, n \in \mathbb{Z}\}$ of even numbers.

- $(P,+)$ is an abelian subgroup of $(\mathbb{Z},+)$.
- Multiplication of even numbers is clearly even: $2n\cdot 2m = 2(2nm)$.
- The distributive rule holds true, as every element in P is also in $\mathbb{Z}$.

Note that $1 \notin P$, therefore there is no identity element for the multiplication in P.

The integers represent a sort of paradigm of rings. However, in general, multiplication can behave rather strangely.

Definition 1.30: In a ring R, two nonzero elements a,b such that $ab = 0$, are (respectively, left and right) **zero divisors**.

Example 1.31: The set $\mathbb{Z}_n$ has a ring structure given by the sum $\oplus$ (see Example 1.25) and a multiplication defined by $\overline{a} \odot \overline{b} = \overline{ab}$. For $n = 4$,

$\odot$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$
$\overline{1}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{2}$	$\overline{0}$	$\overline{2}$	$\overline{0}$	$\overline{2}$
$\overline{3}$	$\overline{0}$	$\overline{3}$	$\overline{2}$	$\overline{1}$

Therefore $\overline{2}$ is a zero divisor, since $\overline{2} \cdot \overline{2} = \overline{4} = \overline{0}$

When an abelian ring with unity $1 \neq 0$ does not have zero divisors, it is called **integral domain**. As we saw, $\mathbb{Z}_4$ is not an integral domain. A particular and very important case of integral domain, is the field.

Definition 1.32: The ring $(R, +, \cdot)$ is a **field** if $(R \setminus \{0\}, \cdot)$ is an abelian group. A generic field will be usually denoted by the symbol $\mathbb{F}$.

Real and complex numbers are the most well known fields. Another very important type of field is given by the construction which was first introduced in Example 1.8 and that will be completed below.

Example 1.33: Using the notation of Example 1.13, we will show here how to equip the abelian group $(Q/\sim, \oplus)$ with a multiplication that turns it into a ring.

As it will become clear soon, $Q/\sim$ is the set of fractions, i.e., the set of rational numbers $\mathbb{Q}$. Therefore, the common symbol $\frac{a}{b} = \overline{(a,b)}$ will be used for its equivalence classes. To further simplify the notation, we will denote the operations simply by '+' and '$\cdot$'.

Therefore, the addition defined in Example 1.13 is

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd},$$

being $0 = \frac{0}{b}$ the identity of the sum, $\forall b \neq 0$. As usual, the additive inverse of $\frac{a}{b}$ is $\frac{-a}{b}$, which will also be written $-\frac{a}{b}$. Moreover, $\frac{a}{b} - \frac{c}{d} = \frac{a}{b} + (-\frac{c}{d})$.

Now we can define multiplication as

$$\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}.$$

Since $\frac{a}{b}\cdot\frac{c}{c}=\frac{ac}{bc}=\frac{a}{b}$, the multiplicative identity is $\frac{1}{1}=\frac{c}{c}$, for any $c\neq 0$, which will be simply denoted by 1.

Finally, the multiplicative inverse of $\frac{a}{b}\neq 0$ is $\frac{b}{a}$, since

$$\frac{a}{b}\cdot\frac{b}{a}=\frac{ab}{ba}=1.$$

Multiplication is also distributive since:

$$\frac{a}{b}\cdot\left(\frac{c}{d}+\frac{e}{f}\right)=\frac{a}{b}\cdot\left(\frac{cf+ed}{df}\right)=\frac{acf+aed}{bdf}=\frac{acbf+aebd}{bdbf}=\frac{a}{b}\cdot\frac{c}{d}+\frac{a}{b}\cdot\frac{e}{f}.$$

In this example, it was shown how to build the field $\mathbb{Q}$ starting from the ring $\mathbb{Z}$. This construction is rather general, as the following theorem states.

Theorem 1.34: The procedure exposed in Example 1.33 can be applied to any integral domain R. The field which is obtained is called **field of fractions** of R.

Definition 1.35: If $(S,+,\cdot)$ is a ring then $R\subseteq S$ is a **subring** (**subfield**) of S if $(R,+,\cdot)$ is a ring (field).

Definition 1.36: Consider two rings $(R,+,\cdot)$ and $(S,\bigoplus,\bigodot)$. The function $f:R\to S$ is an **homomorphism of rings** if, for all $a,b\in R$, the following conditions are verified:

- $f(a+b)=f(a)\bigoplus f(b)$ (it is an homomorphism of the groups $(R,+)$ and $(S,\bigoplus)$).
- $f(a\cdot b)=f(a)\bigodot f(b)$.

When there exists an isomorphism between R and S, they are **isomorphic**, denoted by $R\cong S$.

Remark 1.37: As we did in Remark 1.21, we will say that R is a subring (subfield) of S if R is isomorphic to a subring (subfield) of S, i.e., if there exists an injective ring (field) homomorphism $R\to S$.

Example 1.38: Extending Example 1.22, we can say that $\mathbb{Z}$ is a subring of $\mathbb{Q}$. Indeed, it is easy to prove that $\mathbb{Z} \cong Z = \left\{\frac{a}{1} \in \mathbb{Q}\right\}$ (which is the set introduced in Example 1.22 with the notation of Example 1.33) and that Z is a subring of $\mathbb{Q}$.

Definition 1.39: The characteristic of a ring R with unity, denoted by $\operatorname{char} R$, is the smallest positive p, if it exists, such that, using the notation of Remark 1.15,

$$p \times 1 = 0.$$

If such a p does not exists, we say that R has characteristic 0.

Theorem 1.40: The characteristic of an integral domain must be either 0 or a prime number.

Proof: Consider a ring R with $\operatorname{char} R = p$, $p \neq 0$. We will prove that if p is not a prime number, then R is not an integral domain. So, suppose that $p = mn$, with $m, n \neq 1$. Then,

$$p \times 1 = mn \times 1 = (m \times 1)(n \times 1) = 0.$$

This implies that R has zero divisors, which do not exist in an integer domain. ■

1.5 Vector spaces

Definition 1.41: A **vector space** V over the field $\mathbb{F}$, or $\mathbb{F}$-vector space, is an additive abelian group with an additional operation $(a, v) \in \mathbb{F} \times V \mapsto av \in V$, called **scalar multiplication**, which satisfies the following **linearity conditions**: $\forall a, b \in \mathbb{F}$ and $u, v \in V$,

$$(a+b)v = av + bv \qquad \text{and} \qquad a(u+v) = au + av.$$

Since two types of mathematical objects are involved in vector space operations, the name **vector** will be used for any $v \in V$, while elements $a \in \mathbb{F}$ will be called **scalars**.

A homomorphism f of $\mathbb{F}$-vector spaces, also called $\mathbb{F}$-**linear map**, is a group homomorphism which is compatible with scalar multiplication: if $f : V \to W$, then $f(av) = af(v)$ for every vector v and scalar a.

It is easy to verify that for any vector v in the $\mathbb{F}$-vector space V, $(-a)v = -(a\,v)$, where $-a$ is the additive inverse of $a \in \mathbb{F}$ and $-(a\,v)$ is the additive inverse of $a\,v$ in V. Therefore, the notation $-a\,v$ does not give rise to ambiguities.

Moreover, it is straightforward that $0\,v = 0$, where the 0 on the left is a scalar and the one on the right is a vector. Similarly, $1\,v = v$ holds too.

Definition 1.42: Let V be an $\mathbb{F}$-vector space and $S = \{v_1, \ldots, v_n\} \subseteq V$. The vector $v \in V$ is a **linear combination** of the vector in S with scalars $a_1, \ldots, a_n \in \mathbb{F}$ if

$$v = a_1 v_1 + \cdots + a_n v_n.$$

We also say that v is **generated** by S.

The set S is **linear dependent** if there exist a linear combination equal to the zero vector

$$a_1 v_1 + \cdots + a_n v_n = 0 \tag{1.4}$$

with at least one nonzero scalar. If S is not linearly dependent, it is **linearly independent**.

The meaning of 'linear dependency' is the following: if $a_i \neq 0$ in equation (1.4), then v_i is a **linear combination** of the remaining vectors with coefficients $b_j = -\frac{a_j}{a_i}$, $j \neq i$, i.e.,

$$v_i = b_1 v_1 + \cdots + b_{i-1} v_{i-1} + b_{i+1} v_{i+1} + \cdots + b_n v_n.$$

Moreover, linear independency can be expressed directly as in the following theorem, which is sometimes used as a definition.

Theorem 1.43: The vectors $v_1, \cdots, v_n$ over $\mathbb{F}$ is linearly independent if and only if the equality

$$a_1 v_1 + \cdots + a_n v_n = 0,\ a_i \in \mathbb{F}, i = 1, \ldots, n, \tag{1.5}$$

only holds for $a_1 = a_2 = \cdots = a_n = 0$.

Definition 1.44: Let V be a vector space. The set $B \subseteq V$ is a **basis** of V if it is linear independent and generates (every vector of) V.

If B contains n vectors, then the **dimension** of V is $\dim V = n$.

Theorem 1.45: If V has dimension n then any subset with less than n elements cannot generate V and any subset with more than n elements is linearly dependent.

Definition 1.46: A subset $U \subseteq V$ of an $\mathbb{F}$-vector space V is a **subspace** of V if U itself is an $\mathbb{F}$-vector space, with the same operation.

Theorem 1.47: Given an $\mathbb{F}$-vector space V, $U \subseteq V$ is a subspace of V if and only if $a\,u + b\,v \in U$ for any $u, v \in U$ and $a, b \in \mathbb{F}$

Observe that the vectors generated by any $S \subseteq V$ constitute a subspace of V. This suggests a possible algorithm to construct a basis of a space with dimension n: start from any nonzero vector $v_1 \in V$ and let $B_1 = \{v_1\}$. By Theorem 1.45, it generates a subspace strictly contained in V. Thus, there exists $v_2 \in V$ which is not generated by B_1 and, therefore, $B_2 = v_1, v_2$ is linear independent. Repeating this procedure, the set B_n is obtained, which is a basis of V.

Definition 1.48: We denote by $\text{span}\{v_1, \ldots, v_n\}$ the subspace of V generated by the set of vectors $v_1, \ldots, v_n \in V$.

Definition 1.49: Consider a vector space V with finite dimension. Subspaces U and W which satisfy $V = U + W$ (as in Definition 1.14) are called summands. If for every $v \in V$ there exist unique $u \in U$ and $w \in W$ such that $v = u + w$, then we say that U and W are complementary subspaces, or form a direct sum decomposition of V and write:

$$V = U \oplus W.$$

Theorem 1.50: Let V, U, W be as above. Then, $\dim V = \dim U + \dim W$. Moreover, if B_U B_W are bases of U and V, respectively, then $B_U \cup B_W$ is a basis for V.

Remark 1.51: Since every linearly independent subset of V can be extended to a basis, every subspace has a complement, and the complement is necessarily unique.

Definition 1.52: A symmetric non-degenerate bilinear form on the $\mathbb{F}$-vector space V is the operation $(u,v) \in : V \times V \mapsto u \cdot v \in \mathbb{F}$ which satisfies

- (symmetric) $u \cdot v = v \cdot u$,
- (bilinear) $(a\,u) \cdot v = a(u \cdot v)$, for any $a \in \mathbb{F}$ and
- (non-degenerate) $u \cdot v = 0$ for any $v \in V \iff u = 0$.

Definition 1.53: Let V be a linear space of dimension $n \in \mathbb{N}$ equipped with a symmetric non-degenerate bilinear form and $U \subset V$ a subspace of V of dimension $k \leq n$. The orthogonal complement of U is

$$U^{\perp} = \{v \in V : u \cdot v = 0, \text{for all } u \in U\}.$$

Theorem 1.54: Let V and U are defined as above, with dimension n and k, respectively. Then,

- $U^{\perp}$ is a uniquely determined subspace of V with dimension $n-k$ and
- U and $U^{\perp}$ are complementary, i.e., $V = U \oplus U^{\perp}$.

We end this section establishing some important properties of linear maps.

Definition 1.55: Given the $\mathbb{F}$-linear map $f : V \to W$,

- $\ker_f = \{v \in V : f(v) = 0\}$ is called **kernel** of f.
- $\text{img}_f = \{f(v) : v \in V\}$ is called **image** or range of f.

Theorem 1.56: According notation of Definition 1.55, $\ker_f$ and img_f are subspaces of V and W, respectively.

Definition 1.57: Considering the subspaces defined in Theorem 1.56, we call **nullity** to the dimension of kernel and **rank** to the dimension of **image**.

Theorem 1.58: Considering one $\mathbb{F}$-linear as in Definition 1.55, and denote by m its dimension. We have that

$$\dim \ker_f + \dim \text{img}_f = m.$$

1.6 Matrices and coordinate spaces

Definition 1.59: We will denote by $\mathbb{F}^{p\times q}$ the set of $p\times q$ **matrices**

$$M=\begin{bmatrix} m_{11} & m_{12} & \dots & m_{1q} \\ m_{21} & m_{22} & \dots & m_{2q} \\ \vdots & \vdots & & \vdots \\ m_{p1} & m_{p2} & \dots & m_{pq} \end{bmatrix},$$

where $m_{ij}\in\mathbb{F}$, $i=1,\dots,p$, $j=1,\dots,q$, are called **entries** of M. In a simplified notation, we may also write $M=[m_{ij}]$.

The **transpose** of $M=[m_{ij}]\in\mathbb{F}^{p\times q}$ is $M^T=[m_{ji}]\in\mathbb{F}^{q\times p}$.

When $p=q$ the matrix M is **square**. A square matrix $M=[m_{ij}]$ is **diagonal** if $m_{ij}=0$ whenever $i\neq j$. In this case, we will write $M=\operatorname{diag}(m_{11},m_{22},\dots,m_{pp})$.

The **zero** $p\times q$ matrix, denoted by $0_{p\times q}$, has all entries equal to zero, while the **identity** matrix of dimension p is the $p\times p$ square matrix $I_p=\operatorname{diag}(1,1,\dots,1)$. In both cases, the dimension may be omitted when it is clear from the context.

Theorem 1.60: The set $\mathbb{F}^{p\times q}$ is a $\mathbb{F}$-vector space when equipped with the componentwise operations which are defined as follows: for any $M,N\in\mathbb{F}^{p\times q}$ and $a\in\mathbb{F}$,

- $M+N=[m_{ij}]+[n_{ij}]=[m_{ij}+n_{ij}]=\begin{bmatrix} m_{11}+n_{11} & \cdots & m_{1q}+n_{1q} \\ \vdots & & \vdots \\ m_{p1}+n_{p1} & \cdots & m_{pq}+n_{pq} \end{bmatrix}$ and
- $aM=a[m_{ij}]=[am_{ij}]=\begin{bmatrix} am_{11} & \cdots & am_{1q} \\ \vdots & & \vdots \\ am_{p1} & \cdots & am_{pq} \end{bmatrix}$.

Its dimension is $\dim\mathbb{F}^{p\times q}=pq$.

When one of the dimensions of a matrix is equal to one, we obtain the most simples examples of vector spaces.

Definition 1.61: The vector spaces $\mathbb{F}^{1\times n}$ and $\mathbb{F}^{n\times 1}$ are called **coordinate spaces** and the entries of their vectors are also called **coordinates**.

A vector $v \in \mathbb{F}^{1\times n}$ is called **row vector** and $v \in \mathbb{F}^{n\times 1}$ is called **column vector**. Since we will mainly use column vectors, they will be simply called vectors and we will write $\mathbb{F}^n = \mathbb{F}^{n\times 1}$.

Note that a row vector is easily obtained from a (column) vector by transposition. Actually, if $v \in \mathbb{F}^n$, then $v^T \in \mathbb{F}^{1\times n}$.

Moreover, any matrix $M \in \mathbb{F}^{p\times q}$ can be seen as a sequence of p row vectors of dimension q, called **rows**, or of q column vectors of dimension p, called **columns**.

This different way of looking at a matrix is very useful in order to write linear combinations in a more compact way, as the following result states.

Theorem 1.62: Consider the matrix $M \in \mathbb{F}^{p\times q}$ and the vector $v \in \mathbb{F}^q$. If $M_i \in \mathbb{F}^p$ are the columns of M and v_i the entries of $v, i = 1,\dots,q$, then

$$Mv = M_1 v_1 + \cdots + M_q v_q \in \mathbb{F}^q.$$

In other words, the product Mv is the linear combination of the columns of M whose coefficients are the entries of v.

Theorem 1.63: The **canonical basis** of $\mathbb{F}^n$ is given by the vectors

$$e_1^n = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix},\ e_2^n = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix},\dots, e_n^n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \in \mathbb{F}^n.$$

The canonical basis of $\mathbb{F}^{p\times q}$ is given by matrices $e_i^p (e_j^q)^T \in \mathbb{F}^{p\times q}$, $i = 1,\dots,p$, $j = 1,\dots,q$.

Once again, we will omit the indication of the vector space dimension in its canonical basis vectors when it is fixed or clear from the context, i.e., we will just write $e_1,\dots,e_n$.

To conclude, observe that it is possible to define a product between matrices $(A, B) \in \mathbb{F}^{p\times q} \times \mathbb{F}^{q\times m} \mapsto AB \in \mathbb{F}^{p\times m}$ as follows:

$$AB = [a_{ij}][b_{hk}] = C = [c_{st}],$$

where,

$$c_{ik} = \sum_j a_{ij} b_{jk}. \tag{1.6}$$

The identity matrix I has the following property: for any matrix $A\mathbb{F}^{p\times q}$, $I_p A = A$ and $AI_q = A$. Since the dimension of I is determined by the dimensions of A, we will write only

$$IA = A \qquad \text{and} \qquad AI = A.$$

When $m = 1$, by equation 1.6, the matrix $A \in \mathbb{F}^{p\times q}$ induces an $\mathbb{F}$-linear map $f_A : v \in \mathbb{F}^q \mapsto f_A(v) = Av \in \mathbb{F}^p$.

Definition 1.64: The **kernel**, **image** and **rank** of matrix A are the kernel, image and, respectively, rank of the induced homomorphism f_A.

Theorem 1.65: For any matrix $A \in \mathbb{F}^{p\times q}$, $\mathrm{rk}\, A$ is the maximum number of linear independent columns or rows of A.

Finally, the set of square matrices $\mathbb{F}^{p\times p}$, with the sum and product of matrices we defined, is a ring, as we show in the following example.

Example 1.66: $(R^{2\times 2}, +, \cdot)$ is a ring since it is a group under addition, with identity 0, and multiplication is closed and has identity I. However $(R^{2\times 2}, +, \cdot)$ is not an integral domain, since there exist zero divisors. For example:

$$\begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

1.7 Polynomials

Let R be a subring of S and $X \subseteq S$. We will denote by $R[X]$ the set of all finite linear combinations of powers of elements in X with coefficients in R, i.e.,

$$r(x_1,\ldots,x_n) = \sum_{\text{finite}} r_{i_1,\ldots,i_n} x_1^{i_1} \cdots x_n^{i_n}, \tag{1.7}$$

where, for any $n \in \mathbb{N}$, $i_1,\ldots,i_n \in \mathbb{N}$, $r_{i_1,\ldots,i_n} \in R$ and $x_1,\ldots,x_n \in X$. When $r(x_1,\ldots,x_n) = 0$ if and only if all its coefficients are zero, we say that $x_1,\ldots,x_n \in S$ are **transcendent** over R (otherwise, they are **algebraic**). Note that this is equivalent to say that all the powers of $x_1,\ldots,x_n$ are linear independent over R.

When X is finite and transcendent over R, $R[X]$ is called **ring of polynomials** over R.

Definition 1.67: Given a ring R (with unity), we define the set of (univariate) **polynomials** over R as

$$R[x] = \left\{ r(x) = \sum_{i=0}^{n} r_i x^i,\ r_i \in R, i = 1,\ldots,n, n \in \mathbb{N} \right\}. \tag{1.8}$$

where x is a transcendent element over R, called **variable**, and r_i are called **coefficients**. As for functions, the polynomial $r(x) \in R[x]$ will be often denoted just by r. When all the coefficients of r are zero, it is called **zero polynomial** and we write $r = 0$.

When $r \neq 0$, its **degree**, $\deg r$, is the exponent of its highest power of x with non zero coefficient. If $r = 0$, $\deg r = -\infty$ by convention (being $-\infty + n = n + -\infty = -\infty$ for every $n \in \mathbb{N}$ or $n = -\infty$).

The polynomial r is **constant** when $\deg r < 1$ and **linear** if $\deg r = 1$.

When the coefficient of the highest power of r is equal to 1, r is called **monic**.

Remark 1.68: Note that for any $r \in R[x]$ with $\deg r = n < m$ we can always write $r(x) = \sum_{i=0}^{n} r_i x^i = \sum_{i=0}^{m} r_i x^i$, considering coefficients $r_{n+1} = \cdots = r_m = 0$.

On the set of polynomials $R[x]$ we can define sum and product as follows: given $a(x) =$

$\sum_{i=0}^{n} a_i x^i$ and $b(x) = \sum_{i=0}^{m} b_i x^i$, let

$$a(x)+b(x) = \sum_{i=0}^{\max(m,n)} (a_i + b_i)x^i, \qquad a(x)b(x) = \sum_{i=0}^{n+m} x^i \sum_{j=0}^{i} a_j b_{i-j}. \tag{1.9}$$

Theorem 1.69: The set $R[x]$ with sum and product defined by (1.9) is a ring called **ring of polynomials** with coefficients in R.

If R is commutative, then also $R[x]$ is commutative.

Remark 1.70: Note that by its definition, if $r \in R[x]$, then x is transcendental over R. However, the notation $r(a)$ is well defined also for any $a \in R$, as in (1.7). $r(a)$ is called evaluation of polynomial r at a.

Definition 1.71: A nonconstant polynomial $f(x) \in R[x]$ is **irreducible** over $\mathbb{R}$ if whenever $f(x) = g(x)h(x)$, with $g(x), h(x) \in R[x]$, either $g(x)$ or $h(x)$ is constant.

Theorem 1.72: If $\mathbb{F}$ is a field, for any $r(x) \in \mathbb{F}[x]$ there exist $n \in \mathbb{N}$ unique irreducible polynomials $f_1, \ldots, f_n \in R[x]$ and exponents $\nu_1, \ldots, \nu_n$ such that

$$r(x) = \prod_{i=1}^{n} f_i^{\nu_i}(x).$$

This is the **factorization** of r, f_i are the irreducible **factors**, and ν_i the corresponding **multiplicities**.

Definition 1.73: The **roots** of a polynomials $f \in R[x]$, are the **solutions** of the equation $f(x) = 0$, i.e., the values $a \in \mathbb{F}$ such that $f(a) = 0$.

1.8 Ideals and quotient rings

Definition 1.74: Consider the set $I \subseteq R$, where R is a ring. We say that I is an **ideal** if it satisfies the following conditions (see Definition 1.14 for the notation):

- $IR = RI = I$ and
- $I - I = I$.

Note that, by Theorem 1.20, the second condition means that I is an additive subgroup of R.

Let R be a ring and $I \subseteq R$ an ideal. Now consider the relation on R

$$a \sim b \Leftrightarrow a - b \in I. \tag{1.10}$$

Theorem 1.75: Formula (1.10) defines an equivalence relation.

Proof: The three properties of equivalence relations are satisfied by the additive group structure of I. Indeed,

- $a \sim a$: $a - a = 0 \in I$, because 0 is the identity of I;
- $a \sim b \Rightarrow b \sim a$: if $a - b \in I$, then $b - a = -(a - b) \in I$, since I contains every inverse;
- $a \sim b$ and $b \sim c \Rightarrow a \sim c$: if $a - b \in I$ and $b - c \in I$, also $a - c = (a - b) + (b - c) \in I$, being I closed with respect to the sum.

■

The set $R/\sim$ of equivalence classes defined by (1.10), will be denoted by R/I and the equivalence classes by $\overline{a} = a + I$, for any $a \in R$. As we prove in the following theorem, this set has a 'natural' ring structure.

Theorem 1.76: Let R be a ring and $I \subseteq R$ an ideal. Then $(R/I, +, \cdot)$ is a ring, called **quotient ring**, where the operations are defined by

$$\overline{a} + \overline{b} = \overline{a + b} \quad \text{and} \quad \overline{a} \cdot \overline{b} = \overline{ab}, \quad \forall a, b \in R.$$

In other words, the operations of R are simply extended to set (the equivalence classes) in R/I as in Definition 1.14.

Proof: The demonstration is straightforward once we prove that the operations are well defined. Actually, the ring properties of R/I are just a consequence of the structure of R. For instance, $\overline{0}$ is the additive identity and $\overline{-a}$ the additive inverse of $\overline{a}$.

So, using the notation $\overline{a} = a + I$ and the properties of ideals given in Definition 1.74, observe that, for every $a, b \in R$,

$$\overline{a} + \overline{b} = a + I + b + I = a + b + I = \overline{a + b},$$

where $I + I = I$, since $(I, +)$ is a group, closed with respect to '$+$'. Moreover,

$$\overline{a} \cdot \overline{b} = (a + I)(b + I) = ab + aI + Ib + II = ab + I + I + I = ab + I = \overline{ab},$$

where $aI = Ib = II = I$ are special cases of the defining property $IR = RI = I$. ∎

Example 1.77: We already introduced quotient rings: actually, the ring $\mathbb{Z}_n$ of Example 1.31 is the quotient ring $\mathbb{Z}/n\mathbb{Z}$, where the ideal $n\mathbb{Z}$ is the additive group containing all multiples of n.

Definition 1.78: An ideal I of a ring R is **principal** if $I = aR$ for some $a \in I$, i.e., every element in I is a 'multiple' of a, which is called **generator** of I. The ideal generated by a will also be denoted by (a).

The ring R is a **principal ideal domain**, or PID, if it is an integral domain and every ideal in R is principal.

Theorem 1.79: The ring $\mathbb{Z}$ and any ring of polynomials over a field $\mathbb{F}[x]$ are PIDs.

By the previous theorem, also in the case of polynomial ideals, any ideal is equal to $\big(g\big) = \{a(x)g(x) : a(x) \in \mathbb{F}[x]\}$ for some polynomial $g \in \mathbb{F}[x]$.

Remark 1.80: When no ambiguity exists, we may drop the bar that denotes the equivalence classes in R/I, thus using some special representative. For instance, instead of $\mathbb{Z}_n$ we may use just the set of numbers $\{0, \ldots, n-1\}$ with operations *modulo* n, i.e., the special representative is always the remainder of division by n.

Similarly, in a the polynomial quotient ring $\mathbb{F}[x]/\big(g\big)$ the equivalence class $\overline{a(x)} = a(x) + \big(g\big)$ will be denoted by the remainder of the division of a by g. In other words, $\mathbb{F}[x]/\big(g\big)$ will be represented by all the polynomials of degree less that $\deg g$.

Example 1.81: Let $R = \mathbb{R}[x]/\big(x^2-3\big)$. Then, $\overline{x}$ will be represented by x, while we will choose 3 as a representative of $\overline{x^2}$ and 9 as a representative of $\overline{x^4}$, since $x^2 = 1\cdot(x^2-3)+3$ and $x^4 = (x^2+3)(x^2-3)+9$.

In general, a quotient ring of a PID needs not to be a PID. However, some interesting properties still hold, as we state in the theorem that follow.

Theorem 1.82: Let $g \in \mathbb{F}[x]$ and be I an ideal of the quotient ring $\mathbb{F}[x]/\big(g\big)$. Then,

- I is a principal ideal;
- If $I \neq \mathbb{F}[x]/\big(g\big)$, then (the representative of) its generator divides g.

See [19, Chapter 4.7] for the proof.

Chapter 2

Finite Fields

Such we refereed above, finite fields will have a special importance in this work. In this chapter we present its description and construction, as well as some important theorems and definitions. A **Finite (or Galois) Field** is a field with q elements and will be denoted by $\mathbb{F}_q$. Frequently, in the literature, it is denoted by $\mathbb{F}_q$.

The order of $\mathbb{F}_q$ is q, which is the order of the correspondent additive group. Also $\mathbb{F}_q \backslash \{0\}$ has a multiplicative structure whose order is $q-1$.

2.1 Finite fields with prime order

In this section we will show that the ring $\mathbb{Z}_n$, introduced in Example 1.31, is a field if and only if n is a prime number. For the proof we will need some results about basic properties of integer numbers.

Definition 2.1: Given two numbers $a, b \in \mathbb{Z}$, we define the **greatest common divisor** of a and b by

$$\gcd(a,b) = \max\{c \in \mathbb{N} : c|a \text{ and } c|b\}.$$

If $c = 1$, then a, b are **coprime**.

A famous method to calculate the gcd is the Euclidian algorithm. It works as follows:

Data: a, b
$b_0 = a$;
$b_1 = b$;
$n = 0$;
repeat
  $n \leftarrow n+1$;
  divide b_{n-1} by b_n, with quotient q_n and remainder b_{n+1}: $b_{n-1} = b_n q_n + b_{n+1}$
until $b_{n+1} = 0$;
Result: $\gcd(a,b) = b_n$

Algorithm 1: Euclidean Algorithm

The same algorithm can be used to prove the following important result.

Theorem 2.2 (Bézout equation): Let $a, b \in \mathbb{Z}$ and let $c = \gcd(a,b)$. Then there exist $x, y \in \mathbb{Z}$ such that

$$a\,x + b\,y = c. \tag{2.1}$$

We will prove the theorem constructing x and y. If we let $b_0 = a$ and $b_1 = b$, the equations of Algorithm 2.1 are the following (on the right we write a corresponding matricial equation):

$$\begin{aligned}
b_0 = b_1 q_1 + b_2 \quad &\longleftrightarrow \quad \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = \begin{bmatrix} q_1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \\
b_1 = b_2 q_2 + b_3 \quad &\longleftrightarrow \quad \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} q_2 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} b_2 \\ b_3 \end{bmatrix} \\
&\vdots \\
b_{n-2} = b_{n-1} q_{n-1} + b_n \quad &\longleftrightarrow \quad \begin{bmatrix} b_{n-2} \\ b_{n-1} \end{bmatrix} = \begin{bmatrix} q_{n-1} & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} b_{n-1} \\ b_n \end{bmatrix} \\
b_{n-1} = b_n q_n \quad &\longleftrightarrow \quad \begin{bmatrix} b_{n-1} \\ b_n \end{bmatrix} = \begin{bmatrix} q_n & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} b_n \\ 0 \end{bmatrix}
\end{aligned}$$

Where $b_n = c$. Let $Q_i = \begin{bmatrix} q_i & 1 \\ 1 & 0 \end{bmatrix}$ and observe that

$$\begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = Q_1 \cdots Q_n \begin{bmatrix} b_n \\ 0 \end{bmatrix}. \tag{2.2}$$

If we define $P_i = \begin{bmatrix} 0 & 1 \\ 1 & -q_i \end{bmatrix}$, it is easy to check that $P_i Q_i = I$, for all $i = 1, \ldots, n$. Thus, equation (2.2) can be rewritten as follows:

$$P_n \cdots P_1 \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = \begin{bmatrix} b_n \\ 0 \end{bmatrix}$$

So, if $P_n \cdots P_1 = \begin{bmatrix} x & y \\ s & t \end{bmatrix}$ then

$$\begin{bmatrix} x & y \\ s & t \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = \begin{bmatrix} b_n \\ 0 \end{bmatrix} \quad \Leftrightarrow \quad \begin{bmatrix} x & y \\ s & t \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} c \\ 0 \end{bmatrix} \quad \Leftrightarrow \quad \begin{cases} ax + by = c \\ as + bt = 0 \end{cases}$$

Theorem 2.3: The ring $(\mathbb{Z}_p, +, \cdot)$ is a field if and only if p is a prime.

Proof: '$\Rightarrow$': If $(\mathbb{Z}_p, +, \cdot)$ is a field then it is also an integral domain. By Theorem 1.40 its characteristic is zero or a prime number. Since it is finite, then p is prime.

'$\Leftarrow$': Define $\mathbb{Z}_p^* = \mathbb{Z}_p \setminus \{0\}$ and be p a prime number. We have already seen that $(\mathbb{Z}_p, +)$ is a group $\forall p \in \mathbb{N}$, so we just need to prove that $(\mathbb{Z}_p^*, \cdot)$ is an abelian group. Since multiplication is commutative, it is enough to prove that every element of $\overline{a} \in \mathbb{Z}_p^*$ has a multiplicative inverse, i.e., equation $\overline{a}\,\overline{x} = \overline{1}$ has a solution in $\mathbb{Z}_p^*$. Being $\overline{a} \neq \overline{0}$, a is not a multiple of p, i.e., $a \neq kp$, $\forall k \in \mathbb{Z}$. Thus, once p is prime, $\gcd(a, p) = 1$. By Theorem 2.2, we know that there exist $x, y \in \mathbb{Z}$ such that

$$ax + py = 1 \Rightarrow \overline{ax + py} = \overline{ax} + \overline{py} = \overline{a}\,\overline{x} + \overline{0} = \overline{1}$$

and therefore, $\overline{a}\,\overline{x} = \overline{1}$. ■

Consider $\mathbb{F}_5 = \{0,1,2,3,4\}$ The **Cayley** table for the operations $+$ and $\cdot$ for elements in $\mathbb{F}_5$ are the following:

$+$	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

$\cdot$	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

To conclude this section we give a characterization of zero divisors in $\mathbb{Z}_n$, when it is not a field.

Corollary 2.4: When $\mathbb{Z}_n$ is not an integer domain, the zero divisors are precisely those elements that are not coprime with n.

Proof: Let us consider a nonzero element a of $\mathbb{Z}_n$ which is not coprime with n and be d the greatest common divisor of a and n. Then, $a(\frac{n}{d})$ is equal to $(\frac{a}{d})n$, which is a multiple of n and consequently 0 in $\mathbb{Z}_n$. Thus we have found a nonzero element $b = \frac{n}{d} \in \mathbb{Z}_n$ such that $ab = 0$. Consider now a nonzero element $a \in \mathbb{Z}_n$ which is coprime with n. The existence of $x \in \mathbb{Z}_n$ such that $ax = 1$ can be proved as in Theorem 2.3. Now suppose that there exist $b \in \mathbb{Z}_p$ such that $ab = 0$. Then

$$b = b \cdot 1 = bax = abx = 0$$

Since $b, x \neq 0$, a must be zero. Therefore a is not an zero divisor. ∎

Example 2.5: Consider $a = 6 \in \mathbb{Z}_{16}$. So $d = \gcd(6,16) = 2$. We have that

$$6 \cdot \frac{16}{2} = \frac{6}{2} \cdot 16 = 48 = 0 \text{ in } \mathbb{Z}_{16}.$$

So $a = 6$ and $b = \frac{16}{2} = 8$ are zero divisors.

We already saw that $\mathbb{Z}_p$ is a field is p is prime, now we will see which is, in general, the finite field order.

Theorem 2.6: If $\mathbb{F}$ is a finite field, then it has order p^n for some prime p and $n \in \mathbb{N}$.

Proof: If $\mathbb{F}$ is finite, then $\text{char}\mathbb{F} = p$ with p prime. Also we have that $f : \mathbb{Z}_p \to \mathbb{K} = n \times 1, n = 0, \ldots, p-1 \subseteq \mathbb{F}, n \mapsto n \times 1$ is an isomorphism of fields. Actually, it is an isomorphism of additive groups and then it is easy to check that

$$f(nm) = f(n) \cdot f(m).$$

This means that $\mathbb{F}$ is an extension field of $\mathbb{F}_p$, and therefore is a vector space. And if $\mathbb{F}$ is finite, then it has one finite basis, i.e., $\dim = n$. Hence, an n-dimensional vector space over $\mathbb{F}_q$ with has p^n elements. ■

2.2 Extension Fields

Definition 2.7: Given a field $\mathbb{F}$, we denote by $\mathbb{F}(x)$ the field of fractions of the polynomial ring in one indeterminate $\mathbb{F}[x]$ (which is an integral domain).

Theorem 2.8: The field $\mathbb{F}(x)$ is the smallest field containing $\mathbb{F}$ and x.

Definition 2.9: We say that $\mathbb{F}$ is an **extension** field of $\mathbb{E}$ if $\mathbb{E}$ is a subfield of $\mathbb{F}$ and this relation will be denoted by $\mathbb{F} : \mathbb{E}$. We say that $\mathbb{F}$ is a **simple extension** of $\mathbb{E}$ if $\mathbb{F} = \mathbb{E}(\alpha)$ for some $\alpha \in \mathbb{F}$, called **primitive** element (or generator) of $\mathbb{F}$.

Remark 2.10: We will use the word primitive in a wider sense: if $\mathbb{F} = \mathbb{E}(\alpha)$ we will call primitive every $\beta \in \mathbb{F}$ having the same multiplicative order of $\alpha \in \mathbb{F} \backslash \{0\}$.

From Definition 2.9 it is easy to see that $\mathbb{F}$ is an $\mathbb{E}$-vector space, and we will denote by $[\mathbb{F} : \mathbb{E}]$ its dimension.

Definition 2.11 (Splitting Field): An extension field $\mathbb{F}$ of a field $\mathbb{E}$ is a **splitting field** of a nonconstant polynomial $f(x) \in \mathbb{E}[x]$ if f can be factored into linear factors over $\mathbb{F}$, but not in any proper subfield of $\mathbb{F}$.

Theorem 2.12: Let $\mathbb{E}$ be a finite field and $\mathbb{F}$ one extension field of $\mathbb{E}$. Then $\mathbb{E}$ and $\mathbb{F}$ have the same characteristic.

Theorem 2.13: Let $\mathbb{F} = \mathbb{E}(\alpha)$ where α is algebraic over $\mathbb{F}$. Then

- $\mathbb{F} = \mathbb{E}[\alpha]$;
- As a $\mathbb{E}$-vector space, $\dim \mathbb{F} = n < \infty$, being $1, \alpha, \ldots, \alpha^{n-1}$ a basis of $\mathbb{F}$ over $\mathbb{E}$;
- The minimal degree of $r \in \mathbb{E}[x]$ such that $r(\alpha) = 0$ is n.

Remark 2.14: In the conditions of the previous theorem, the extension field can be represented by the set of polynomials in $\mathbb{E}[x]$ with degree less than n or, equivalently, by $\mathbb{E}^n$. In particular, the isomorphisms can be chosen so that:

$$a \in \mathbb{F} \quad \longleftrightarrow \quad b(x) = b_0 + b_1 x + \cdots + b_{n-1} x^{n-1} \in \mathbb{E}[x] \quad \longleftrightarrow \quad c = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_{n-1} \end{bmatrix}.$$

where $a = b(\alpha)$.

Observe that only the additive structure of $\mathbb{F}$ is taken into account considering these vector space isomorphisms: the multiplicative structure of the field will be investigated in the following sections.

Example 2.15: Consider a field $\mathbb{Q}$ and a polynomial $f(x) = x^2 - 3 \in \mathbb{Q}[x]$, and let ξ be a root of f. Clearly ξ it is not an element of $\mathbb{Q}$ but f is one polynomial with coefficients in $\mathbb{Q}$, therefore ξ is an algebraic element over $\mathbb{Q}$. Now consider the set $\mathbb{Q}[\xi]$. We will show that any $a \in \mathbb{Q}[\xi]$ can be written as a linear polynomial in ξ and therefore $\mathbb{Q}[\xi]$ is an extension field of $\mathbb{Q}$. Indeed, suppose that $a = \sum_{i=0}^{M} a_i \xi^i$ with M odd (otherwise add a zero coefficient). Then

$$a = \sum_{i=0}^{M} a_i \xi^i = \sum_{i=0}^{\frac{M-1}{2}} a_{2i} \xi^{2i} + \sum_{i=0}^{\frac{M-1}{2}} a_{2i+1} \xi^{2i+1} = \sum_{i=0}^{\frac{M-1}{2}} a_{2i} 3^i + \sum_{i=0}^{\frac{M-1}{2}} a_{2i+1} 3^i \xi = \alpha_0 + \alpha_1 \xi,$$

where $\alpha_0, \alpha_1 \in \mathbb{Q}$. Now we can show that $\mathbb{Q}[\xi]$ is a field, i.e, any nonzero element $a \in \mathbb{Q}[\xi]$ has an inverse. In other words, we will prove that the equation $ab = 1$ has a solution $b \in \mathbb{Q}[\xi]$. Considering $a = \alpha_0 + \alpha_1\xi \in Q[\xi]$ and $b = \beta_0 + \beta_1\xi \in Q[\xi]$,

$$ab = \alpha_0\beta_0 + (\alpha_0\beta_1 + \alpha_1\beta_0)\xi + (\alpha_1\beta_1)\xi^2 = \alpha_0\beta_0 + 3\alpha_1\beta_1 + (\alpha_0\beta_1 + \alpha_1\beta_0)\xi = 1 \Leftrightarrow$$
$$\begin{bmatrix} \alpha_1 & \alpha_0 \\ \alpha_0 & 3\alpha_1 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

It is well known that this linear system has a (unique) solution if and only if the determinant of the coefficients matrix is not zero, i.e,

$$3\alpha_1^2 - \alpha_0^2 \neq 0 \Leftrightarrow 3\alpha_1^2 \neq \alpha_0^2 \Leftrightarrow \alpha_0 \neq \pm\sqrt{3}\alpha_1,$$

which is true $\forall \alpha_0, \alpha_1 \in \mathbb{Q}$.
Note that $\mathbb{Q}[\xi]$ is the set of polynomials with degree less than 2 which is equivalent to say that the $\mathbb{Q}$-vector space $\mathbb{Q}[\xi]$ has dimension 2 and $\{1, \xi\}$ is a basis for $\mathbb{Q}[\xi]$. When we write $a = \alpha_0 + \alpha_1\xi \longleftrightarrow [\alpha_0\ \alpha_1]$ we are using the vectorial structure of the previous remark.

2.3 Multiplicative structure of finite fields

What that we demonstrated in Section 2.1 can be repeated replacing $\mathbb{Z}$ by $\mathbb{F}[x]$, because the properties we used are just based on the concepts of factors and division, which are almost equal in both rings.

Theorem 2.16: The ring $\mathbb{F}_p[x]/\big(g\big)$ is a field if and only if $g \in \mathbb{F}_p[x]$ is irreducible.

Proof: Suppose that $\mathbb{F}_p[x]/\big(g\big)$ is not irreducible. Then there exist factors h and k, such that $g = hk$, corresponding to $\overline{h} \cdot \overline{k} = \overline{0}$, analogue to the proof of Theorem 1.40. ■

Corollary 2.17: Consider $g \in \mathbb{F}_p[x]$ irreducible with $\deg g = n$. Then $\mathbb{F}_p[x]/\big(g\big) \cong \mathbb{F}_{p^n}$.

Theorem 2.18: For each $\beta \in \mathbb{F}_q$ there exists only one monic polynomial $p(x)$ of minimal degree in $\mathbb{F}_q[x]$ such that:

- $p(\beta)=0$.
- The degree of $p(x)\leqslant m$.
- If there is a polynomial $f(x)\in\mathbb{F}_q[x]$ such that $f(\beta)=0$ then $p(x)|f(x)$.
- $p(x)$ is irreducible in $\mathbb{F}_q[x]$.

Remark 2.19: In Corollary 2.17 g is the generator polynomial of $\mathbb{F}_q$. If g_1 and g_2 are irreducibles with $\deg n$, then

$$\mathbb{F}_q[x]/\big(g_1\big)\cong\mathbb{F}_q[x]/\big(g_2\big)\cong\mathbb{F}_{p^n}.$$

However the multiplicative structure is different, as we will show in example 2.23.

Theorem 2.20: Every elements of $\mathbb{F}_q$ satisfy the equation

$$x^q-x=0. \tag{2.3}$$

Furthermore, they constitute the entire set of roots of this equation.

Proof: We can rewrite the equation 2.3 as follows

$$x^q-x=0\Longleftrightarrow x(x^{q-1}-1)=0$$

Hence 0 is clearly a root. The nonzero elements of the field are all generated as powers of α, being α an primitive element of the $\mathbb{F}_q$. So $\forall\beta\in\mathbb{F}_q\backslash\{0\},\exists i\in\mathbb{Z}:\beta=\alpha^i$ and

$$\beta^{q-1}=(\alpha^i)^{q-1}=(\alpha^{q-1})^i=1^i=1.$$

Since there are q elements in $\mathbb{F}_q$, and q roots for the equation, the elements of $\mathbb{F}_q$ are all the roots. ■

Theorem 2.21: An element $\beta\in\mathbb{F}_{p^m}$ lies in $\mathbb{F}_p$ if and only if $\beta^p=\beta$.

Proof: By theorem 2.20, $\forall \beta \in \mathbb{F}_p \backslash \{0\}$, $\beta^p = \beta$. Conversely, assume that $\beta = \beta^p$.Then β is a root of $x^p - x = 0$. All p elements of $\mathbb{F}_p$ satisfy this polynomial and it only has p roots. As $\beta \in \mathbb{F}_{p^m}$, it lies that β satisfies the equation $\beta^{p^m} = \beta, \forall m \geq 0$. ■

Example 2.22: The field $\mathbb{F}_{64}$ is an extensive field of $\mathbb{F}_4$. Let α be a primitive in $\mathbb{F}_{64}$. We look for an element of $\mathbb{F}_{64}$ which is also element of $\mathbb{F}_4$. Consider $\beta = \alpha^{21}$. So

$$\beta^4 = \alpha^{21 \cdot 4} = \alpha^{63}\alpha^{21} = \beta$$

Then, by theorem 2.21, $\beta \in \mathbb{F}_4$ and therefore $\mathbb{F}_4 = \{0, 1, \beta, \beta^2\} = \{0, 1, \alpha^{21}, \alpha^{42}\}$.

Example 2.23: Consider the polynomials $g_1(x) = x^3 + x + 1$ and $g_2(x) = x^3 + x^2 + 1$ over $\mathbb{F}_8$.

Let say that α is a primitive element of $\mathbb{F}_{2^3}$ over $\mathbb{F}_2$, such that α is a root of $g_1(x)$, i.e.,

$$\alpha^3 = \alpha + 1. \tag{2.4}$$

Now take successive powers of a beyond α^3:

$$\begin{aligned}
&\alpha^3 = 1 + \alpha^2, \\
&\alpha^4 = \alpha \cdot (\alpha^3) = \alpha + \alpha^2, \\
&\alpha^5 = \alpha \cdot (\alpha^4) = \alpha^2 + \alpha^3 = 1 + \alpha + \alpha^2, \\
&\alpha^6 = \alpha^2 \cdot (\alpha^4) = \alpha^3 + \alpha^4 = 1 + \alpha + \alpha + \alpha^2 = 1 + \alpha^2, \\
&\alpha^7 = \alpha^6 \cdot (\alpha) = \alpha + \alpha^3 = 1 \\
&\quad \vdots
\end{aligned}$$

(Note that, by Theorem 2.12, $\mathrm{char}\,\mathbb{F}_{2^3} = \mathrm{char}\,\mathbb{F}_2 = 2$. Thus $\alpha^i + \alpha^i = 0$.)
With linear combinations of powers of α with maximum degree 2, we represented powers of α up to α^7 all distinct, and $\alpha^7 = 1$ because this is a cyclic group. This is called **power representation** of the elements of a group.

Using Theorem 2.13, we can define an element $\beta \in \mathbb{F}_8$ such that β is a linear combination of powers of α, as follows:

$$\beta = a + b\alpha + c\alpha^2 \in \mathbb{F}_2[\alpha].$$

According Remark 2.14, we can write β in a vector representation form: $\beta = [a\ b\ c]^T \in \mathbb{F}_2^3$.

In table **??** we depicted the relation between vector representation, polynomial representation and power representation of the elements of $\mathbb{F}_8$ generated by the polynomial $g_1(x)$, and in table2.1 we represent the same relation but when $\mathbb{F}_8$ is generated by the polynomial $g_2(x)$. Analyzing those tables, it is possible to verify the isomorphism presented in Remark 2.17.

Power Representation	Polynomial Representation	Vector Representation	Numeric Representation
0	0	000	0
1	1	001	1
α	α	010	2
α^2	α^2	100	4
α^3	$\alpha^2 + 1$	101	5
α^4	$\alpha^2 + \alpha + 1$	111	7
α^5	$+\alpha + 1$	011	3
α^6	$\alpha^2 + \alpha$	110	6

Table 2.1: $\mathbb{F}_{2^3}$ generated by $p(x) = x^3 + x^2 + 1$

Example 2.24: Let exemplify the algorithm of multiplication between two elements of $\mathbb{F}_{2^3}$ when they are in the power representation. (The direct correspondence between the different representations of the elements of the field will be denoted by $\longleftrightarrow$, and are in table **??**).

Let α be one primitive element of $\mathbb{F}_{2^3}$ and $g_1 = x^3 + x + 1$ the respective primitive polynomial. Now consider α^6 and α^4 elements of $\mathbb{F}_{2^3}$. Using the polynomial representation we have that:

$$\alpha^6 \cdot \alpha^4 = (1+\alpha^2)\cdot(\alpha+\alpha^2) = \alpha + \alpha^2 + \alpha^3 + \alpha^4 = \alpha + \alpha^2 + 1 + \alpha + \alpha + \alpha^2 = 1 + \alpha = \alpha^3.$$

Note that:

$$\alpha^6 \cdot \alpha^4 = \alpha^{10 \mod 7} = \alpha^3 \longleftrightarrow [1\ 1\ 0]^T.$$

Generalizing, the multiplication of elements when they are in its power representation

are made modulo $q-1$, i.e.,

$$\alpha^m \cdot \alpha^n = \alpha^{(m+n) \bmod (q-1)}. \tag{2.5}$$

Power Representation	Polynomial Representation	Vector Representation	Numeric Representation
0	0	0000	0
1	1	0001	1
α	α	0010	2
α^2	α^2	0100	4
α^3	α^3	1000	8
α^4	$\alpha+1$	0011	3
α^5	$\alpha^2+\alpha$	0110	6
α^6	$\alpha^3+\alpha^2$	1100	12
α^7	$\alpha^3+\alpha+1$	1011	11
α^8	α^2+1	0101	5
α^9	$\alpha^3+\alpha$	1010	10
α^{10}	$\alpha^2+\alpha+1$	0111	7
α^{11}	$\alpha^3+\alpha^2+\alpha$	1110	14
α^{12}	$\alpha^3+\alpha^2+\alpha+1$	1111	15
α^{13}	$\alpha^3+\alpha^2+1$	1101	13
α^{14}	α^3+1	1001	9

Table 2.2: $\mathbb{F}_{2^4}$ generated by $p(x)=x^4+x+1$

2.4 Minimal and primitive polynomials

Definition 2.25: Let $\beta \in \mathbb{F}_{q^m}$. The **minimal** polynomial of β with respect to $\mathbb{F}_q$ is the smallest-degree, nonzero, monic polynomial $p(x) \in \mathbb{F}_q[x]$ such that $p(\beta)=0$.

Definition 2.26: Let $\beta \in \mathbb{F}_{q^m}$. The **conjugates** of β with respect to the subfield $\mathbb{F}_q$ are $\beta, \beta^q, \beta^{q^2}, \beta^{q^3}, \cdots$ and form a set called the **conjugacy class** of β with respect to $\mathbb{F}_q$.

Power Representation	Polynomial Representation	Vector Representation	Numeric Representation
0	0	00	0
1	1	01	1
α	α	10	2
α^2	$\alpha+1$	11	3

Table 2.3: $\mathbb{F}_{2^2}$ generated by $p(x)=x^2+x+1$

Example 2.27: Let $\alpha \in \mathbb{F}_{2^4}$ be a primitive element. The conjugates of α with respect to $\mathbb{F}_2$ are:

$$\alpha, \alpha^2, \alpha^{2^2}=\alpha^4, \alpha^{2^3}=\alpha^8, \alpha^{2^4}=\alpha$$

So the conjugacy class of α is

$$(\alpha, \alpha^2, \alpha^4, \alpha^8).$$

Let $\beta=\alpha^3$ be an element which is not in conjugacy class of α. The conjugacy class of of β is:

$$\beta=\alpha^3, (\alpha^3)^2=\alpha^6, (\alpha^3)^{2^2}=\alpha^{12}, (\alpha^3)^{2^3}=\alpha^9, (\alpha^3)^{2^4}=\alpha^3$$

So the conjugacy class of β is

$$(\alpha^3, \alpha^6, \alpha^9, \alpha^{12}).$$

Taking now $\gamma=\alpha^5$, (another unused element), its conjugacy class is:

$$\gamma=\alpha^5, (\alpha^5)^2=\alpha^{10}, (\alpha^5)^{2^2}=\alpha^5$$

So the conjugacy class of γ is (α^5, α^{10}).

Continuing taking the next unused element, let $\delta=\alpha^7$; Its conjugacy classe is:

$$\delta=\alpha^7, (\alpha^7)^2=\alpha^{14}, (\alpha^7)^{2^2}=\alpha^{13}, (\alpha^7)^{2^3}=\alpha^{11}, (\alpha^7)^{2^4}=\alpha^7$$

The conjugacy class of δ is consequently $(\alpha^7, \alpha^{11}, \alpha^{13}, \alpha^{14})$.

The other elements of $\mathbb{F}_{2^4}$ are 1 and 0, which always forms its own conjugacy classes. The conjugacy classes are thus partitions of the field.

Theorem 2.28: Let $\beta \in \mathbb{F}_{q^m}$ with $\text{ord}(\beta) = n$, and let d be the multiplicative order of q modulo n,i.e., the smallest positive integer such that $n|q^d - 1$. Then $\beta^{q^d} = \beta$. The elements $\beta, \beta^q, \beta^{q^2}, \cdots, \beta^{q^{d-1}}$ are all distinct.

Theorem 2.29: Let $\beta \in \mathbb{F}_{q^m}$ have order n and let d be the multiplicative order of $q \mod n$. Then $p(x) = \prod_{i=0}^{d-1}(x - \beta^{q^i}) \in \mathbb{F}_q[x]$. Furthermore, $p(x)$ is irreducible, which means that $p(x)$ is the **minimal** polynomial for β.

Note that for each $\beta \in \mathbb{F}_{q^m}$ of order n, and being d be the multiplicative order of q modulo n, the set of $\beta^{q^i}, i = 1, \ldots, n-1$ is one conjugacy class of $\mathbb{F}_{q^m}$. In this sense, the conjugacy classes over $\mathbb{F}_{2^3}$ with respect to $\mathbb{F}_2$ for example, have the following minimal polynomials:

- $0 \rightarrow M(x) = x$.
- $1 \rightarrow M(x) = x + 1$.
- $\alpha, \alpha^2, \alpha^4 \rightarrow M(x) = (x - \alpha)(x - \alpha^2)(x - \alpha^4) = x^3 + x + 1$
- $\alpha^3, \alpha^6, \alpha^5 \rightarrow M(x) = (x - \alpha^3)(x - \alpha^5)(x - \alpha^6) = x^3 + x^2 + 1$.

Definition 2.30: (Primitive Polynomial) An irreducible polynomial $p(x) \in \mathbb{F}_p[x]$ of degree m is said to be **primitive** if the smallest positive integer n for which $p(x)$ divides $x^n - 1$ is $n = p^m - 1$.

Theorem 2.31: The roots of an m-th degree primitive polynomial $p(x) \in \mathbb{F}_p[x]$ are primitive elements in $\mathbb{F}_{p^m}$. Every primitive polynomial is the minimal polynomial of some primitive element.

Example 2.32: Consider the polynomial $f(x) = x^4 + x + 1$. By exhaustive search we can see that $f(x) \not| x - 1, \cdots, f(x) \not| x^{13} - 1$, and $f(x) \not| x^{14} - 1$ but $f(x) | x^{15} - 1$. Actually

$$x^{15} - 1 = (x^4 + x + 1) \cdot (x^{11} + x^8 + x^7 + x^5 + 2x^4 + x^3 + x^2 + 3x + 3).$$

$15 = 2^4 - 1$ so $f(x)$ is a primitive polynomial of $\mathbb{F}_{2^4}$.

2.5 Galois Field Fourier Transform

Is possible to define a Fourier transform over a sequence of Galois field numbers.

Definition 2.33: Let $v = (v_0, v_1, \cdots, v_{n-1})$ be a vector over $\mathbb{F}$ of length n such that $n|q^m - 1$ for some $m \in \mathbb{N}$. Let $\alpha \in \mathbb{F}_{q^m}$ have order n. The **Galois field fourier transform** (GFFT) of v is the vector $V = (V_0, V_1, \cdots, V_{n-1})$ with components

$$V_j = \sum_{i=0}^{n-1} \alpha^{ij} v_i \;\; j = 0, 1, \cdots, n-1 \tag{2.6}$$

We write $V = \mathfrak{F}[v]$ and $v \longleftrightarrow V$ to denote the Fourier transform relationship between v and V.

Definition 2.34: In a field $\mathbb{F}$ with characteristic p, we call **Inverse Galois Field Fourier Transform** of the vector $V = (V_0, V_1, \cdots, V_{n-1})$ to a vector v defined as follows

$$v_i = n^{-1} \sum_{j=0}^{n-1} \alpha^{-ij} V_j \tag{2.7}$$

where n^{-1} is the multiplicative inverse of n modulo p.

Definition 2.35: The **spectrum** of the polynomial $v(x) = v_0 + v_1 x + \cdots, + v_{n-1} x^{n-1}$ is the GFFT of $v = (v_0, v_1, \cdots, v_{n-1})$.

Chapter 3

The Behavioral Approach to Systems Theory

This chapter contains the key definitions of this thesis. We start by defining a dynamical system and a mathematical model, and then we explain the behavioral approach as it was described by J. C. Willems in[30].

3.1 Dynamical System

We start by introducing the behavior approach in a dynamical system.

Definition 3.1: A **dynamical system** Σ is defined as a triple

$$\Sigma = (\mathbb{T}, \mathbb{W}, \mathfrak{B}), \tag{3.1}$$

where $\mathfrak{B}$ is the **behavior** and represent a set of functions, called **trajectories**, having domain $\mathbb{T}$ and codomain $\mathbb{W}$, i.e., $\mathfrak{B} \subseteq \{w : \mathbb{T} \to \mathbb{W}\} = \mathbb{W}^{\mathbb{T}}$.

In this work we will deal with discrete-time systems where $\mathbb{T} = \mathbb{Z}$ and $\mathbb{W} = \mathbb{F}^q$.

Definition 3.2: A dynamical system $\Sigma = (\mathbb{T}, \mathbb{W}, \mathfrak{B})$ is said to be **linear** if $\mathbb{W}$ is a vector space over $\mathbb{F}$ and $\mathfrak{B}$ is a linear subspace of $\mathbb{W}^{\mathbb{T}}$, i.e.,

$$\forall\, w_1, w_2 \in \mathfrak{B} \text{ and } \forall \alpha, \beta \in \mathbb{F}, \alpha w_1 + \beta w_2 \in \mathfrak{B}$$

3.2 Polynomial Matrices and Operators

We use the same notation as in Section 1.6 to represent **polynomial matrices**,

$$R(s) = [r_{ik}(s)] \in \mathbb{F}^{g\times q}[s], \tag{3.2}$$

i.e., matrices whose entries are polynomials. Notice that this set is not a vector space over $\mathbb{F}[s]$, since this is not a field. However,

- the operations defined in Section 1.6 (sum, product and transposition) still hold for polynomial matrices;
- linear independence of polynomial vectors, over $\mathbb{F}[s]$ (and therefore the rank of polynomial matrices) can not be defined as in Section 1.6, since that definition is based on the invertibility of the coefficients.

To overcome the latter difficulty observe that we may always consider polynomials as a subring of their field of fraction, $\mathbb{F}[s] \subset \mathbb{F}(s)$. So, the polynomial matrix (3.2) belongs to the $\mathbb{F}(s)$-vector space $\mathbb{F}^{g\times q}(s)$.

Definition 3.3: A set of polynomial vectors is said to be linearly independent over $\mathbb{F}[s]$ if it is linearly independent over $\mathbb{F}(s)$. The rank of a polynomial matrix is the number of its linearly independent (polynomial) rows or columns.

Theorem 3.4: The polynomial vectors $v_1(s),\ldots,v_n(s)$ are linearly independent if and only if the equality

$$a_1(s)v_1(s)+\cdots+a_n(s)v_n(s)=0, \tag{3.3}$$

where $a_i(s)\in\mathbb{F}[s]$, is satisfied only by $a_1(s)=\cdots=a_n(s)=0$.

Proof: By Theorem 1.43, the vectors are linearly independent if and only if equation (3.3), with $a_i(s)\in\mathbb{F}(s)$, is satisfied only by $a_1(s)=\cdots=a_n(s)=0$. So, we have to show that this condition does not change if $a_i(s)$, $i=1,\ldots,n$ are polynomials. In particular, we prove the theorem by showing that if condition (3.3) can only be satisfied by null polynomials then it can be satisfied only by null fractions.

Actually, suppose that the condition holds for fractions $a_1(s),\ldots,a_n(s)$ being at least one of them not zero. It multiply all the fractions by the least common multiple of their denominators, $d(s)$, then $d(s)a_i(s)\in\mathbb{F}[s]$. Therefore,

$$d(s)a_1(s)v_1(s)+\cdots+d(s)a_n(s)v_n(s)=d(s)\cdot 0=0,$$

is a nontrivial solution of equation (3.3) with polynomial coefficient. ■

Definition 3.5: Let $R(s)\in\mathbb{F}^{g\times q}(s)$ be a polynomial matrix and denote the rows of $R(s)$ by $r_i(s)$, $i=1,\ldots,g$. The **row degrees** $d_1,\ldots,d_g$ are defined as $d_i=\max\deg r_{ij}(s)$, $j=1,\ldots,q$.

Definition 3.6: Consider a polynomial $R(s)\in\mathbb{F}^{g\times q}[s]$ such that $R(s)=\frac{N(s)}{D(s)}$. If $\deg D(s)>\deg N(s)$ then $R(S)$ is **strictly proper**. If $N(s)$ and $D(S)$ have the same degree, $R(s)$ is **proper**.

Definition 3.7: Let $n_1,\ldots,n_q$ be nonnegative integers and consider the polynomial matrices $R(s),N(\lambda)\in\mathbb{F}^{q\times q}(s)$ such that

$$N(\lambda)=\operatorname{diag}(\lambda^{n_1},\ldots,\lambda^{n^q}).$$

The row degrees of the matrix $R(s)N(\lambda)$ are the **weighted row degrees** of $R(s)$.

Definition 3.8: The square polynomial matrix $R(s)\in\mathbb{F}^{g\times g}(s)$ **unimodular** if it admits a polynomial inverse, i.e., a polynomial matrix $Q(s)\in\mathbb{F}^{g\times g}(s)$ such that $R(s)Q(s)=Q(s)R(s)=I$.

3.3 Autoregressive Models (AR)

Consider the following matrix difference equation in the trajectory w:

$$R_l\,w(t+l)+R_{l-1}\,w(t+l-1)+\cdots+R_0\,w(t)=0, \tag{3.4}$$

where $R_0,\cdots,R_l\in\mathbb{R}^{g\times q}$. This equation is called **auto regressive** (AR) model. We can write Equation (3.4) more compactly using the following operator.

Definition 3.9: Consider a function $w : \mathbb{Z} \to \mathbb{F}^n$. We define the **shift operator** σ, acting on elements w, defined as follows

$$(\sigma w)(k) = w(k+1). \tag{3.5}$$

Indeed, since $\sigma^i w(t) = w(t+i)$, equation (3.4) equal to write:

$$\begin{aligned} R_e\sigma^e w(t) + r_{e-1}\sigma^{e-1} w(t) + \cdots + R_1\sigma w(t) + R_0 w(t) &= 0 \\ (R_e\sigma^e + r_{e-1}\sigma^{e-1} + \cdots + R_1\sigma + R_0)w(t) &= 0 \\ R(\sigma)w(t) &= 0 \end{aligned} \tag{3.6}$$

where we are using a polynomial matrix operator obtained, formally, from the polynomial matrix

$$R(s) = R_e\sigma^e + r_{e-1}\sigma^{e-1} + \cdots + R_1\sigma + R_0$$

This equation defines a dynamic system in sense of definition 3.1 as:

$$\Sigma(R) = (\mathbb{T}, \mathbb{W}, \mathfrak{B}),$$

where the behavior $\mathfrak{B}$ is defined as

$$\mathfrak{B} = \{w : \mathbb{T} \to \mathbb{W}, \text{ such that } R(\sigma)w(t) = 0 \text{ for all } t \in \mathbb{T}\}. \tag{3.7}$$

The polynomial matrix $R(s)$ is a **kernel** representation of the behavior $\mathfrak{B} = \ker R$. Note that different representations of a matrix $R(s)$ may define the same behavior. The following definition qualify the set of matrices which define the same behavior.

Definition 3.10: Two matrices $R_1(s)$ and $R_2(s)$ are said to be **equivalent** if they define the same behavior.

Theorem 3.11: Given two matrices $R_1(s)$ and $R_2(s)$, there exist $U(s), V(s)$ such that $R_2(s) = U(s)R_1(s)$ and $R_1(s) = V(s)R_2(s)$.

Corollary 3.12: Let $R(s)$ be a polynomial matrix. Then $\ker R = \ker UR$ for any unimodular matrix $U(s)$.

Definition 3.13: For all $w \in \mathbb{W}$, $\forall A \subseteq \mathbb{T}$, $w_{|A} : A \to \mathbb{W}$, such that $w_{|A}(t) = w(t), \forall\, t \in \mathbb{T}$.

Definition 3.14: A dynamical system $\Sigma = (\mathbb{T}, \mathbb{W}, \mathfrak{B})$, with $\mathbb{T} = \mathbb{Z}$, is said to be **time-invariant** if $\sigma\mathfrak{B} = \mathfrak{B}$.

Definition 3.15: A dynamical system $\Sigma = (\mathbb{T}, \mathbb{W}, \mathfrak{B})$ is said to be **complete** if $\{w_{|[t_1,t_2]} \in \mathfrak{B}_{|[t_1,t_2]} : \forall\, t_1, t_2 \in \mathbb{T}, t_1 \leq t_2\}$.

Theorem 3.16: Let $\mathbb{T} = \mathbb{Z}$ and $\mathbb{W} = \mathbb{R}^q$ and consider the system $\Sigma = (\mathbb{T}, \mathbb{W}, \mathfrak{B})$. Then there exists a polynomial matrix $R(s)$ such that $\mathfrak{B} = \ker R$ if and only if Σ is linear, time invariant, and complete.

Actually, the behavior $\mathfrak{B}$ defined in equation 3.7 is linear and time invariant. Consider w_1 and $w_2 \in \mathfrak{B}$. We have that it is linear, because

$$R(\sigma)[\alpha w_1 + \beta w_2] = R(\sigma)\alpha w_1 + R(\sigma)\beta w_2 = \alpha R(\sigma)w_1 + \beta R(\sigma)w_2 = 0 \tag{3.8}$$

And time-invariant, since

$$R(\sigma)(\sigma w_1) = \sigma R(\sigma)w_1 = 0, \tag{3.9}$$

$\mathfrak{B}$ is also time invariant.

3.4 Mathematical Models

The propose of this section is to expose an approach to mathematic models. The framework depicted leads in the following direction: find algorithms for obtain models from observed data. We will start establishing a mathematical language. Suppose that we have a particular phenomenon that we want to model. We define a set S, called **universum**, which contains all elements produced in the phenomenon. The elements of S are called **attributes** of the phenomenon.

Naturally, when we model a phenomenon, some attributes are constant. Thus, we define a subset $M \subseteq S$, which excludes the attributes that are not considered, and this is the **model**.

This means that the model will only produce outcomes in M. To represent the family of models, we define the subset $\mathfrak{M} \subseteq 2^S$. The measurements will be in a subset $Z \subseteq S$, that can be thought as a experimental evidence summarizing. Also here we define $\mathfrak{Z} \subseteq 2^S$ which is the class of measurement sets.

We say that a model M is **unfalsified** by the measurements Z if $Z \subseteq M$. However, given a set Z, there are a wide number of models M compatibles with. The question arises in which one to choose. We say that a model M_1 is more **powerful** than M_2 if $M_1 \subseteq M_2$.

Definition 3.17: We say that M_Z^* is the **most powerful unfalsified model**(MPUM) in the model class $\mathfrak{M}$ based on the measurements Z if $Z \subseteq M_Z^* \in \mathfrak{M}$ and $Z \subseteq M \in \mathfrak{M}, M_Z^* \subseteq M$.

Theorem 3.18: Take $S = V$ and $\mathfrak{M} = \{$ all linear subspaces of $V\}$. Then $M_Z^* = \text{span}\{z | z \in Z\}$.

3.5 The MPUM for Dynamical Systems

A dynamic system is a particular case of a mathematical system in which the phenomenon produces outcomes that are functions of time and thus, the universum is a function space. In the notation of Section 3.4, $S = \mathbb{W}^{\mathbb{T}}$ is the vector space of all possible trajectories (having infinite dimension), $\mathfrak{M}$ is the set of all subspaces of S with finite dimension, and Z is a set of trajectories.

Given a set of trajectories Z, to find the MPUM is equal to find a polynomial matrix $R^*(s)$, which is the kernel of the minimal behavior $\mathfrak{B}$ such that $Z \subset \mathfrak{B}$, i.e., $\mathfrak{B}$ is the vector space generated by the trajectories in Z.

Theorem 3.19: Let $w : \mathbb{Z} \to \mathbb{F}^q$. Then there exists a most powerful AR model given by aware representation $R(s)$:

$$R(\sigma) w_i = 0, \forall i.$$

Given a set of trajectories, the algorithm to find the MPUM is the following

Data: $\{w_1,\dots,w_n\}\subseteq(\mathbb{F}^q)^{\mathbb{Z}}$
$R_0(s)=I$;
for i=*1 to* n **do**
  $\Delta_i = R_{i-1}(\sigma)w_i$ (define the $i-th$ error trajectory);
  $V_i(\sigma)\Delta_i = 0$ (compute a kernel representation of the MPUM for the trajectory w_i);
  $R_i(s) = V_i(s)R_{i-1}(s)$ (compute the MPUM for the set $\{w_1,\dots,w_i\}$;
end
Result: Kernel representation of the MPUM: $R_n(s)$

Algorithm 2: MPUM General Algorithm

At the beginning we assume that our behavior is a null space, say $\mathfrak{B}_0=\{0\}$. Equivalently, $R_0(s)=I$ and therefore $\dim\ker R_0=0$ and $\operatorname{rk}R_0=q$.

In the first step, we include the first observed trajectory w_1 and $\mathfrak{B}_1=\operatorname{span}\{w_1\}$. Thus $\dim\mathfrak{B}_1=1$ and we find the matrix $R_1(s)$ so that $R_1(s)$ is the kernel representation of $\mathfrak{B}_1$ and $\operatorname{rk}R_1(s)$ is minimal.

Then we include the second trajectory w_2 and $\mathfrak{B}_2=\operatorname{span}\{w_1,w_2\}$. We check if w_2 is in the null space $R_1(s)$. If yes, $R_2(s)$ is equal to $R_1(s)$ and follow including the next trajectory; If not, we calculate $R_2(S)$ so that $R_2(S)$ is the kernel representation of $\mathfrak{B}_2$ and $\dim\ker R_2=\dim\ker R_1+1$ which, according Theorem 1.58, implies that and $\operatorname{rk}R_2=\operatorname{rk}R_1-1$.

And we proceed in such a way until we have included all observed trajectories. At the end we obtain the matrix $R_n(s)$ with minimum rank, such that $R_n(S)$ it is the kernel representation of $\mathfrak{B}_n$, the vector space generated for all trajectories $w_1,\dots,w_n$, i.e., the MPUM.

We will consider only trajectories, (measurements), of the form $w_i(k)=b_i\lambda_i^k,\ b_i\in\mathbb{F}^q$, because are satisfied the conditions of the following theorem:

Theorem 3.20: Let $R(s)\in\mathbb{R}^{q\times q}[s]$, and $\det R(s)$ be a polynomial of degree n, and let

$$\mathfrak{B}=\{w:\mathbb{Z}\to\mathbb{F}^q \mid R(\sigma)w=0\}.$$

If the roots of $\det R(s)$ are mutually distinct and belong to $\mathbb{F}$, say $\det R(s)=\prod_{i=1}^n(\xi-\lambda_i)$, with $\lambda_i\in\mathbb{F}$, then all trajectories in $\mathfrak{B}$ are of the form

$$w(k)=\sum_{i=1}^{n}b_i\lambda_i^k,$$

with $b_i \in \mathbb{F}^q$ such that $R(\lambda_i)b_i = 0$.

This algorithm will be applied later in Chapter 5.

Chapter 4

Codes

Coding Theory study methods for an efficient and accurate transfer of information from one device to another. The physical medium through which the information is transmitted is called a channel. It happens that undesirable disturbances, such as **noise**, may affect the information in the transmitted process along the channel, and therefore errors in the received data may occur. Create robust codes that avoid such disturbances is the aim of coding theory. In this chapter we analyze the linear codes, particularly the class of cyclic linear codes, and we show the construction of the BCH and the RS codes.

4.1 Linear Block Codes

Linear codes are the most well-studied and practically used codes. They have strong structural property, which provides guidance in the search of new good codes, and have a practical encoder and decoder process. Although, if we want to obtain the largest possible number of codewords, we must sometimes use nonlinear codes. In this work we are only concerned in linear codes.

Definition 4.1: A **code** C is a non-empty subset of $\mathbb{F}^n$, where $\mathbb{F}$ is a field called **alphabet**. If $\mathbb{F}$ is a finite field and C is a vector space over $\mathbb{F}$, then C is a **linear block code**. The **length** of the code is n and the **dimension** of the code is $\dim C$. Any element of C is called **codeword** or block, and the number of codewords is the **size** of C. An (n,k) linear code is any linear code of length n and dimension k.

Remark 4.2: Note that, by definition of vector space, a linear code C is translation-invariant, i.e.

$$c \in C \Rightarrow c + C = C$$

Definition 4.3: We denote by $d(c_1, c_2)$ the **Hamming distance** between two codewords c_1 and c_2, which is equal to the number of nonzero entries of $c_1 - c_2$. For all $c_1, c_2 \in C \subseteq \mathbb{F}^n$, $0 \leq d(c_1, c_2) \leq n$.

Remark 4.4: Actually, the hamming distance satisfies the properties of distance:

- $0 \leq d(c_1, c_2) \leq n$,
- $d(c_1, c_2) = 0 \iff c_1 = c_2$,
- $d(c_1, c_2) = d(c_2, c_1)$,
- (Triangular Inequality) $d(c_1, c_2) \leq d(c_1, c_2) + d(c_2, c_3)$,

for all $c_1, c_2 \in \mathbb{F}^n$.

Definition 4.5: The **minimum distance** of a code C is:

$$d_{\min}(C) = \min\{d(a, b) : a, b \in C, a \neq b\}. \tag{4.1}$$

Remark 4.6: Observe that the minimum distance of a codeword must be at least 1.

Definition 4.7: The **weight** of a codeword $c \in C$ is equal to the number of nonzero entries in the vector. The **minimum weight** of a code C is the smallest weight of any non-zero code word of C: $w_{\min}(C) = \min\{w(c) : c \in C\}$.

Theorem 4.8: For any code C, the minimum distance and minimum weight coincide, i.e.,

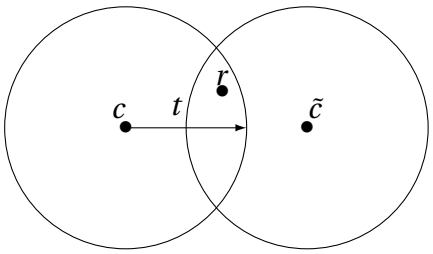
$$d_{\min}(C) = w_{\min}(C). \tag{4.2}$$

Proof: For all $c \in C, w(c) = d(c,0) \Rightarrow \exists c \in C : w_{\min}(C) = w(c) = d(c,0) \geq d_{\min}(C)$.

Conversely, there exist $c_1, c_2 \in C$ such that $d_{\min}(C) = d(c_1, c_2) = d(c_1 - c_2, 0) = w(c_1 - c_2) \geq w_{\min}(C)$. Thus $d_{\min}(C) = w_{\min}(C)$.

Note that we used linearity when we assumed that $c_1 - c_2$ is a codeword. ■

Suppose that a code word is transmitted and a single error is made by the channel. Then the Hamming distance from the received word to the transmitted codeword is equal to 1. Now suppose that t errors occur during transmission. The errors can be **detected** if and only if $d_{\min}(C) \geq t+1$. Consider the following picture which satisfies the previous condition

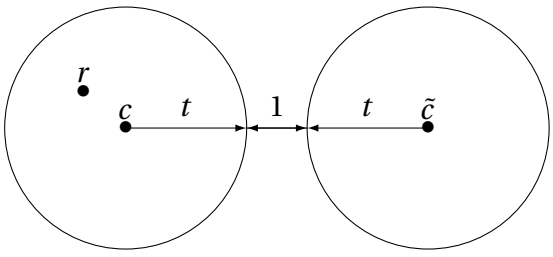


where c and $\tilde{c}$ are respectively the transmitted word and an other word of C, r is the received word, and every words $\tilde{r}$ inside the circle satisfy the condition $d(\tilde{r}, \bar{c}) \leq t$, where $\bar{c}$ correspond to the codeword of the center of the circle. For every r with errors, $d(r, \tilde{c}) \neq 0$ and $d(r, c) \neq 0$ and thus r is not a code word; this is the condition for error detection.

When we detect that r has errors, we correct the errors by associating r to the codeword of the middle of the center in which it is inside. However, regarding the previous picture, we cannot understand if r correspond to c or $\tilde{c}$. Therefore, we can **correct** t errors if

$$d_{\min}(C) \geq 2t+1. \tag{4.3}$$

Proof: Consider the following picture:



Let say that $d(r,\tilde{c}) = d_2$ and $d(r,c) = d_1$, $(d_1 \leq t)$. By the triangle inequality we know that:

$$2t+1 \leq d(c,\tilde{c}) \leq d(c,r) + d(r,\tilde{c}) \leq t + d_2.$$

Thus,

$$t + d_2 \geq 2t+1 \Leftrightarrow d_2 \geq t+1 \geq d_1,$$

and therefore we will associate r to c.

■

Theorem 4.9: The minimum distance for an (n,k) linear code is bounded by

$$d_{\min} \leq n-k+1.$$

4.1.1 Matrix description of Linear Block Codes

Linear codes can be studied using linear algebra theory and tools.

Consider a linear code $C \in \mathbb{F}^n$ generated by vectors $g_1,\ldots,g_m$.Then, by Theorem 1.62, if g_i are columns of matrix G,

$$C = \operatorname{img} G \tag{4.4}$$

i.e., for any $c \in C$, there exist $a \in \mathbb{F}^m$ such that $c = Ga$.

Remark 4.10: If C is an (n,k) code, then there exist $G \in \mathbb{F}^{n\times k}$ satisfying (4.4). Indeed, in this case, the columns of G are a basis of C.

Definition 4.11: Consider a code $C \subseteq \mathbb{F}_q^n$ of dimension k. By Corolary1.54, there exists an orthogonal complement $C^\perp$ of dimension $n-k$. By definition $C^\perp$ is also a code, and is called **dual code** of C.

Theorem 4.12: Denote by H a matrix whose columns generate $C^\perp$. Then

$$H^T c = 0$$

Theorem 4.12 can be used to check whether a word is or not a codeword of C. Therefore H is a **check matrix** of the code C.

Theorem 4.13 ([19, Theorem.3.3]): Let H be a parity check matrix of the linear block code C. Then $d_{\min}(C)$ is equal to the smallest positive number of rows of H which are linearly dependent.

4.1.2 Cyclic Codes

Cyclic codes are a subclass of linear codes. We have seen that a linear code over $\mathbb{F}_q$ can be described in terms of a check matrix H.

Definition 4.14: A linear code C over $\mathbb{F}_q$ is called **cyclic code** if it is invariant under a cyclic shift:

$$(c_0, c_1, \ldots, c_{n-1}) \in C \Rightarrow (c_{n-1}, c_0, \ldots, c_{n-2}) \in C \tag{4.5}$$

Lemma 4.15: Let $c(x) \longleftrightarrow (c_0, \ldots, c_{n-1})$. Then, in $\mathbb{F}(x)/(x^n-1)$, $x\,c(x) \longleftrightarrow (c_{n-1}, c_0, \ldots, c_{n-2})$.

Proof: We will use here the 'bar' notation for elements in $\mathbb{F}(x)/(x^n-1)$. So,

$$\begin{aligned}\overline{x\,c(x)} &= \overline{c_0 x + c_1 x^2 + \cdots + c_{n-1} x^n} = \overline{c_0 x + c_1 x^2 + \cdots + c_{n-2} x^{n-1} + c_{n-1}(x^n - 1) + c_{n-1}} \\ &= \overline{c_{n-1} + c_0 x + c_1 x^2 + \cdots + c_{n-2} x^{n-1}},\end{aligned}$$

being $\overline{c_{n-1}(x^n-1)} = \overline{0}$. ∎

Theorem 4.16: C is a cyclic code if and only if it is an ideal of $\mathbb{F}_q[x]/(x^n-1)$.

Proof:

'⇒' By definition 1.74, C is an ideal if it is a group and if it is closed under multiplication by a scalar. Since C is a vector space, it is verifies those properties and therefore it as an ideal.

'⇐'

Now let C be an ideal of $\mathbb{F}_q[x]/(x^n-1)$ and let $c \in C$. By definition of ideal, if $c \in C$, also $x\,c \in C$, $x \in \mathbb{F}_q[x]/(x^n-1)$. By Lemma 4.15, the multiplication by x correspond to a cyclic

shift of the coefficients of c. Therefore, if c represents the polynomial whose coefficients are the elements of a codeword, xc is still a codeword and thus C is a cyclic code. ■

Remark 4.17: As in Remark 1.80, $\mathbb{F}_q[x]/(x^n-1)$ and its subsets, are represented by a polynomial with degree less than n.

Therefore, by Theorem 1.82, every code C is an ideal generated by some polynomial g, called **generator polynomial** of C, and contains its multiples with degree less than n. If $\deg g = n-k$, then $C = \{p(x)g(x),\ \deg p(x) < k\}$. Therefore, $\dim C = k$. This statement proves the first part of Theorem 1.82.

4.2 BCH Codes

BCH codes are a multiple error correction codes. They were discovered in 1959 by Hocquenghem, and independently in 1960 by Bose and Ray-Chaudhuri. The abbreviation BCH comprises the initials of these inventors' names. One of the key features of BCH codes is that during code design, there is a precise control over the number of symbol errors correctable by the code.

4.2.1 Design of BCH Codes

A BCH code over a $\mathbb{F}_p$ of length n, capable of correcting t errors is specified as follows:

- Determine the smallest m such that $\mathbb{F}_{q^m}$) has a primitive nth root of unity β.
- Select a nonnegative integer b. Usually, $b = 1$.
- Consider $2t$ consecutive powers of β starting from β^b:

$$\beta^b, \beta^{b+1}, \cdots, \beta^{b+2t-1}.$$

  Determine the minimal polynomial with respect to $\mathbb{F}_p$ of each of these powers. (Because of conjugacy, these minimal polynomials need not to be distinct)

- The generator of the code $g(x)$ is the least common multiple (LCM) of these minimal polynomials. The code generated by $g(x)$ is an $(n, n-\deg g)$ cyclic code. Because the generator polynomial is constructed using minimal polynomials with respect to $\mathbb{F}_p$, the generator $g(x)$ has coefficients in $\mathbb{F}_p$, and the code is over $\mathbb{F}_{p^m}$.

If $b=1$ in the construction procedure, the BCH code is said to be **narrow sense**. If $n=q^m-1$ then the BCH code is said to be **primitive**.

In the construction of a BCH code, the generator polynomial has the coefficients in $\mathbb{F}_p$, called the "small field", and the roots in $\mathbb{F}_p^m$, called the "big field". The code words are also in the "small field". The "big field" will be necessary only for de decoding procedure.

Example 4.18: Let $n=63=2^6-1$ for a primitive code with $m=6$ and let β be a root of the primitive polynomial $x^4+x^3+x^2+x+1$ in $\mathbb{F}_2^4$. Suppose that we want a narrow sense ($b=1$) and a triple-error correcting binary BCH code.

So, following the design steps described above, we need $2t=6$ consecutive powers of β: $\beta, \beta^2, \beta^3, \beta^4, \beta^5, \beta^6$. Now, dividing them into conjugacy classes with respect to $GF(2)$ we have:

$$\{\beta, \beta^2, \beta^4\}, \{\beta^3, \beta^6\}, \{\beta^5\}.$$

According2.29, the corresponding minimal polynomials for these sets are:

$$P_1(x)=x^5+x^2+1\,; P_2(x)=x^5+x^4+x^3+x^2+1 \text{ and } P_3(x)=x^5+x^4+x^2+x+1$$

So the generator polynomial $g(x)$ is:

$$\begin{aligned} g(x) &= LCM[P_1(x), P_2(x), P_3(x)] = p_1 \cdot p_2 \cdot p_3 \\ &= x^{15}+x^{11}+x^{10}+x^9+x^8+x^7+x^5+x^3+x^2+x+1; \end{aligned}$$

This gives a (63,63-15)=(63,48) binary cyclic code.

4.3 Reed–Solomon Codes

RS codes are very similar to BCH Codes, with just a few differences. One is the generator polynomial $g(x)$ construction. We saw that for BCH, $g(x)$ is build over $\mathbb{F}_q$, the "small field". In RS codes it is build over $\mathbb{F}_{q^m}$, the "big field". This is because a RS code is a q^m-ary BCH code of length q^m-1. The advantage is that in RS we can choose the exact number of roots of the generator polynomial.

Lemma 4.19: The minimum distance of an (n,k) Reed-Solomon code is $d_{\min} = n-k+1$.

Proof: The polynomial message $c(x)$ has at most $k-1$ roots, since it is a polynomial of degree k. There are at most $k-1$ zero positions in each nonzero codeword. Thus $d_{\min} \geq n-(k-1)$. However, according to theorem 4.9 we must have $d_{\min} \leq n-k+1$. So $d_{\min} = n-k+1$. ■

Considering α a primitive element of $\mathbb{F}_p$ over $\mathbb{F}_{p^m}$. The generator polynomial for a RS code is given by

$$g(x) = (x-\alpha^b)(x-\alpha^{b+1})\cdots(x-\alpha^{b+2t-1}),$$

and $g(x)$ has degree $2t$. Thus $n-k=2t$ for a RS code and the design distance is $\delta = n-k+1$.

4.4 Construction of RS Codes

There are two different ways of constructing RS codes. We will make a description of those constructions and present a connection between them.

4.4.1 First RS Construction

Let α be a primitive element over $\mathbb{F}_{p^m}$ and let $n = p^m-1$. Now consider the message vector $m = (m_0, m_1, \cdots, m_{k-1}) \in \mathbb{F}_{p^m}^k$ and its associated polynomial message $m(x) = m_0 + m_1 x + \cdots + m_{k-1}x^{k-1} \in \mathbb{F}_{p^m}[x]$. The encoding is defined by the mapping $\rho : m(x) \longmapsto c$ by

$$(c_0, c_1, \ldots, c_{n-1}) \triangleq \rho(m(x)) = (m(1), m(\alpha), m(\alpha^2), \cdots, m(\alpha^{n-1})). \tag{4.6}$$

That is, $\rho(m(x))$ evaluates $m(x)$ at all the non-zero elements of $\mathbb{F}_{p^m}$.

Example 4.20: Consider a primitive element α over $\mathbb{F}_{2^2}$ and the respective primitive polynomial $p(x) = x^2 + x + 1$, as in table (2.3). Now suppose that we want to transmit the sequence:

$$m = (2, 1), \tag{4.7}$$

which corresponds to

$$m = (\alpha, 1), \tag{4.8}$$

in $\mathbb{F}_{2^2}$. Therefore the corresponding polynomial matrix is

$$m(x) = \alpha + x. \tag{4.9}$$

According to equation 4.6, the code word is

$$c = (m(1), m(\alpha), m(\alpha^2), m(\alpha^3)) \tag{4.10}$$

where

$$\begin{aligned} m(1) &= \alpha + 1 = \alpha; \\ m(\alpha) &= \alpha + 1 = \alpha^2; \\ m(\alpha^2) &= \alpha + \alpha = 0; \\ m(\alpha^3) &= \alpha + \alpha^2 = 1; \end{aligned}$$

$$c = (\alpha, \alpha^2, 0, 1) \tag{4.11}$$

4.4.2 Second RS Construction

Given a message $m = (m_0, m_1, \ldots, m_{k-1})$ and the corresponding polynomial representation $m(x) = m_0 + m_1 + \ldots + m_{k-1}x^{k-1}$, where $m_i \in \mathbb{F}_q$, the systematic encoding process is

$$c(x) = m(x)x^{n-k} - r(x) \tag{4.12}$$

where $r(x)$ denotes the remainder after division by the generator of C, g. Since c(x) is an ideal,

$$\begin{aligned} c(x) &= m(x)x^{n-k} - r(x) = \\ &= h(x)g(x) + r(x) - r(x) = \\ &= h(x)g(x) \end{aligned} \tag{4.13}$$

Example 4.21: Consider t=2. The consecutive powers of α are

$$\alpha, \alpha^2, \alpha^3, \alpha^4$$

The generator polynomial is

$$g(x) = (x-\alpha)(x-\alpha^2)(x-\alpha^3)(x-\alpha^4) \tag{4.14}$$

4.4.3 Equivalence of the two RS Code Constructions

Theorem 4.22: Let $n|q^m-1$ for some m. A q-ary-n-tuple with weight$\leq \delta-1$ that also has $\delta-1$ consecutive zeros in its spectrum must be the zero vector. That is, the minimum weight of the code is $\geq \delta$.

According to theorem 4.22, a RS code has a consecutive sequence of $2t = d_{\min}-1$ zeros in its GFFT. We know form Lemma 4.19 that the minimum distance of a RS code is $d_{\min} = n-k+1$. We now show that the codewords constructed according construction 1 have a consecutive sequence of $n-k$ zeros in their spectrum. Consider the polynomial

$$m(x) = m_0 + m_1 x + \cdots + m_{k-1}x^{k-1}$$

and let the codeword constructed according equation (4.6) be

$$c = (m(1), m(\alpha), \ldots, m(\alpha^{n-1})) \tag{4.15}$$

so that

$$c_i = m(\alpha^i) = \sum_{l=0}^{k-1} m_l \alpha^{il},\ i = 0, 1, \ldots, n-1. \tag{4.16}$$

Computing the GFFT of c we get

$$C_j = \sum_{i=0}^{n-1} c_i \alpha^{ij}, \tag{4.17}$$

where the index $-j$ is to be interpreted cyclically. Now substituting c_i in equation (4.17) we obtain

$$C_j = \sum_{i=0}^{n-1} \sum_{l=0}^{k-1} m_l \alpha^{-ij} \alpha_{il} = \sum_{l=0}^{k-1} \left[\sum_{i=0}^{n-1} \alpha^{i(l-j)} \right]. \tag{4.18}$$

The inner summation is 0 if $l \neq j \mod n$. This is the case for $-j = k, k+1, \cdots, n-1$, in which is $n-k$ consecutive values of j. Thus, there are $n-k$ consecutive zeros in the GFFT of every codeword.

Chapter 5

Decoding BCH and RS Codes

In this last chapter we present some algorithms to decode BCH codes and RS codes, giving special attention to the application of the "behavioral decoding´´ for RS codes, with an practical example.

5.1 The general outline for decoding BCH and RS Codes

There are many algorithms developed to decode BCH and RS codes. In this section we just present the general outline, and in the next sections we describe concepts and present some algorithms. The general steps to decode RS and BCH codes are the follow

- Computation of the syndrome.
- Determination of an error locator polynomial, whose roots provide an indication of the position errors.
- Finding the roots of the error locator polynomial.
- For RS codes or nonbinary BCH codes, also the error values must be determined.

5.1.1 Syndrome and Error Pattern

Consider $g(x)$ a generator polynomial of a code C and α a primitive element of some field $\mathbb{F}_q$. Since

$$g(\alpha) = g(\alpha^2) = \cdots = g(\alpha^{2t}) = 0$$

and according to equation(4.13), it follows that a codeword $c = (c_0, \ldots, c_{n-1})$ with polynomial $c(c) = c_0 + \cdots + c_{n-1}x^{n-1}$ satisfies

$$c(\alpha) = \cdots = c(\alpha^{2t}) = 0,$$

since $c(x)$ is multiple of $g(x)$. Consider now a channel error

$$e(x) \longleftrightarrow (e_0, e_1, \ldots, e_{n-1}),\ e_j \in \mathbb{F}_q. \tag{5.1}$$

For the received data $r(x) = c(x) + e(x)$ we calculate:

$$S_j = r(\alpha^j) = e(\alpha^j) = \sum_{k=0}^{n-1} e_k \alpha^{jk},\ j = 1, 2, \ldots, 2t. \tag{5.2}$$

Where $S_1, \ldots, S_{2t}$ are called **syndromes** of the received data. Suppose that r has v errors in it which are at locations $i_1, \ldots, i_v$. The error locations are those values of j such that $e_{i_j} \neq 0$.. So, if we let $X_k = \alpha^{i_k}$,

$$S_j = \sum_{k=1}^{v} e_{i_k}(\alpha^j)^{i_k} = \sum_{k=1}^{v} e_{i_k}(\alpha^{i_k})^j = \sum_{k=1}^{v} e_{i_k} X_k^j,\ j = 1, 2, \ldots, 2t. \tag{5.3}$$

If we consider $e_j \in \mathbb{Z}_2$, we can simplify the syndromes, obtaining

$$S_j = \sum_{k=1}^{v} X_k^j,\ j = 1, 2, \ldots, 2t. \tag{5.4}$$

If we know X_k, we can deduce the exact position of the errors. For example, consider $X_j = \alpha^5$. By definition of X_k it means that $i_k = 5$ and thus the error is in the received digit r_5. The X_k are the **error locators**.

5.1.2 The Error Locator Polynomial

There are several different ways of finding the locator polynomial. Equation (5.4) leads to a system with $2t$ equations and v unknowns, (the error locators), which can be solved exhaustively once $2t \geq v$. However this process may be computationally unattractive, thus another technique is adopted. We define the **error locator polynomial** as

$$\Lambda(x) = \prod_{k=1}^{v}(1 - X_k x) = \Lambda_v x^v + \Lambda_{v-1} x^{v-1} + \cdots + \Lambda_1 x + \Lambda_0. \tag{5.5}$$

where $\Lambda_0 = 1$. We can see that the roots of this polynomial are $X_1^{-1}, \ldots, X_v^{-1}$ which are the reciprocals of the error locators.

5.1.3 Chien Search

The goal now is, given an error locator polynomial, to find its roots. This is usually done by Chien search, which is an exhaustive search over every elements in $\mathbb{F}_q$. We just have to examine every elements and check if it is a root. The process presented below is know by **Chien search**.

Example 5.1: Consider thet we have 4 errors($v = 4$). The error locator polynomial is

$$\Lambda(x) = \Lambda_4 x^4 + \Lambda_3 x^3 + \Lambda_2 x^2 + \Lambda_1 x + 1$$

Now we just need to evaluate $\Lambda(x)$ at each nonzero elements in $\mathbb{F}_q$ as follows:

$$\begin{aligned}
\Lambda(1) &= \Lambda_4(1)^4 + \Lambda_3(1)^3 + \Lambda_2(1)^2 + \Lambda_1(1) + 1, \\
\Lambda(\alpha) &= \Lambda_4(\alpha)^4 + \Lambda_3(\alpha)^3 + \Lambda_2(\alpha)^2 + \Lambda_1(\alpha) + 1, \\
&\vdots \\
\Lambda(\alpha^{q^m-2}) &= \Lambda_4(\alpha^{q^m-2})^4 + \Lambda_3(\alpha^{q^m-2})^3 + \Lambda_2(\alpha^{q^m-2})^2 + \Lambda_1(\alpha^{q^m-2}) + 1,
\end{aligned}$$

finding all the roots of Λ.

5.1.4 Finding the Error Locator Polynomial

If we expand equation (5.5) for $\nu = 4$ we obtain:

$$\begin{aligned}\Lambda &= 1 - x(X_1 + X_2 + X_3 + X_4) + x^2(X_1X_2 + X_1X_3 + X_1X_4 + X_2X_3 + X_2X_4 + X_3X_4) \\ &\quad - x^3(X_1X_2X_3 + X_1X_2X_4 + X_1X_3X_4) + x^4X_1X_2X_3X_4 \\ &= \Lambda_0 + x\,y_1 + x^2\Lambda_2 + x^3\Lambda_3 + x^4\Lambda_4\end{aligned}$$

so that

$$\begin{aligned}\Lambda_0 &= 1 \\ \Lambda_1 &= -(X_1 + X_2 + X_3 + X_4) \\ \Lambda_2 &= X_1X_2 + X_1X_3 + X_1X_4 + X_2X_3 + X_2X_4 + X_3X_4 \\ \Lambda_3 &= -(X_1X_2X_3 + X_1X_2X_4 + X_1X_3X_4) \\ \Lambda_4 &= X_1X_2X_3X_4\end{aligned}$$

In general, for an error locator polynomial of degree ν we find that

$$\begin{aligned}&\Lambda_0 = 1, \\ &-\Lambda_1 = \sum_{k=1}^{\nu} X_k = X_1 + X_2 + \cdots + X_\nu, \\ &\Lambda_2 = \sum_{k<m}^{\nu} X_kX_m = X_1X_2 + X_1X_3 + \cdots + X_1X_\nu + \cdots + X_{\nu-1}X_\nu, \\ &-\Lambda_3 = \sum_{k<m<n}^{\nu} X_kX_mX_n = X_1X_2X_3 + X_1X_2X_4 + \cdots + X_{\nu-2}X_{\nu-1}X_\nu, \\ &\vdots \\ &(-1)^\nu\Lambda_\nu = X_1X_2\cdots X_\nu.\end{aligned}$$

As the following theorem states, there is a linear relation between the syndromes and the coefficients of the error locator polynomial.

Theorem 5.2: The syndromes defined by equation (5.4) and the coefficients of the error

locator polynomial defined by equation (5.5), are related by

$$S_k + \Lambda_1 S_{k-1} + \cdots + \Lambda_{k-1} S_1 + k\Lambda_k = 0, 1 \leq k \leq v$$
$$S_k + \Lambda S_{k-1} + \cdots + \Lambda_{v-1} S_{k-v+1} + \Lambda_v S_{k-v} = 0, k > v.$$

From the last theorem we get

$$S_k = -\sum_{i=1}^{v} \Lambda_i S_{k-i},\ k > v. \tag{5.6}$$

The system of equation (5.6) can be expressed in matrix form as follows:

$$\begin{bmatrix} S_1 & S_2 & \cdots & S_v \\ S_2 & S_3 & \cdots & S_{v+1} \\ \vdots & & & \\ S_v & S_{v+1} & \cdots & S_{2v-1} \end{bmatrix} \begin{bmatrix} \Lambda_v \\ \Lambda_{v-1} \\ \vdots \\ \Lambda_1 \end{bmatrix} = - \begin{bmatrix} S_{v+1} \\ S_{v+2} \\ \vdots \\ S_{2v} \end{bmatrix}.$$

Where the $[S_{ij}], i, j = 1, \ldots, v$ is denoted by M_v, is constant on the skew diagonals. Since the number of errors is not known in advance, an algorithm is needed.

5.1.5 Peterson-Gorenstein-Zierler Algorithm

The Peterson-Gorenstein-Zierler algorithm decoder operates as follows,

- Set $v = t$
- Form M_v and determine if M_v is invertible, (for instance, computing its determinant). If it is not invertible, set $v \leftarrow v - 1$ and repeat this step.
- If M_v is invertible, solve for the coefficients $\Lambda_1, \Lambda_2, \ldots, \Lambda_v$.

5.1.6 Berlekamp-Massey Algorithm

The main feature of Berlekamp-Massey algorithm is that at each stage of the algorithm there is the possibility of to re-use information that has already been learned. Equation (5.6) describes the output of a **linear recurrent sequence**(LRS) with coefficients $\Lambda_1, \Lambda_2, \ldots, \Lambda_v$.

For this formula to work properly, we must find the coefficients Λ_i, in such a way that the LRS generates the known sequence of syndromes $S_1, S_2, \ldots, S_{2t}$. In Berlekamp-Massey algorithm we built the LRS that produces the entire sequence $\{S_1, S_2, \ldots, S_{2t}\}$ by successively modifying an existing LRS. We start with an LRS that produces S_1. Then we check if that LRS could also produce the sequence $\{S_1, S_2\}$; if yes, we do not make any alteration, if not we determine a new LRS a longer sequence. Proceeding inductively in this way, we start from an LRS capable of producing the sequence $S_1, S_2, \ldots, S_{k-1}$ and modify it, if necessary, so that it can produce the sequence $S_1, S_2, \ldots, S_k$. At each stage, the modifications to the LRS should be done in such way that the LRS is the shorter possible. At the last stage of the algorithm we should be able to produce the sequence $\{S_1, S_2, \ldots, S_{2t}\}$ and its coefficients correspond to the error locator polynomial $\Lambda(x)$ of smallest degree. Note that this is a special case of application the MPUM, where the S_k are the trajectories and the MPUM for thus trajectories is the error locator polynomial $\Lambda(x)$ of smallest degree. Let L_k denote the length of the LRS produced at stage k of the algorithm and

$$\Lambda^{[k]}(x) = 1 + \Lambda_1^{[k]} x + \cdots, + \Lambda_{L_k}^{[k]} x^{L_k}$$

be the **connection polynomial** at stage k, indicating the connections for the LRS capable of producing the output sequence $\{S_1, S_2, \ldots, S_k\}$,i.e.,

$$S_j = -\sum_{i=1}^{L_k} \Lambda_i^{[k]} S_{j-i}, \ j = L_k + 1, \ldots, k. \tag{5.7}$$

At some intermediate step, suppose we have a connection polynomial

$$\Lambda^{[k-1]}(x) = 1 + \Lambda_1^{[k-1]} x + \cdots, + \Lambda_{L_{k-1}}^{[k-1]} x^{L_{k-1}}$$

of length L_{k-1} that produces the sequence $\{S_1, S_2, \ldots, S_{k-1}\}$ for some $k-1 < 2t$. We check if this connection polynomial also produces S_k by computing the output

$$\widetilde{S}_k = -\sum_{i=1}^{L_{k-1}} \Lambda_i^{[k-1]} S_{k-i}.$$

If $\widetilde{S}_k$ is equal to S_k, then there is no need to update the LRS, so $\Lambda^{[k]}(x) = \Lambda^{[k-1]}(x)$ and

$L_k = L_{k-1}$. Otherwise, there is an error associated with $\Lambda^{[k-1]}(x)$,

$$d_k = S_k - \widetilde{S}_k = S_k + \sum_{i=1}^{L_{k-1}} \Lambda_i^{[k-1]} S_{k-i} = \sum_{i=0}^{L_{k-1}} \Lambda_i^{[k-1]} S_{k-i}. \tag{5.8}$$

In this case, we update the connection polynomial using the formula

$$\Lambda^k(x) = \Lambda^{[k-1]}(x) + Ax^p \Lambda^{[m-1](x)} \tag{5.9}$$

where $A \in \mathbb{F}_q$, p is an integer, and $\Lambda^{[m-1]}(x)$ is one of the prior connection polynomials produced by our process with nonzero discrepancy d_m. Using this new connection polynomial, we compute the new discrepancy, denoted by d'_k, as

$$d'_k = \sum_{i=0}^{L_k} \Lambda_i^{[k]} S_{k-i} = \sum_{i=0}^{L_{k-1}} \Lambda_i^{[k-1]} S_{k-i} + A \sum_{i=0}^{L_{m-1}} \Lambda_i^{[m-1]} S_{k-i-p}. \tag{5.10}$$

Now consider $p = k - m$. Then by comparison with the definition of the discrepancy in 5.8, the second summation gives

$$A \sum_{i=0}^{L_{m-1}} \Lambda_i^{[m-1]} S_{m-i} = A d_m.$$

Thus, if we choose $A = -d_m^{-1} d_k$, then the summation in equation 5.10 gives

$$d'_k = d_k - d_m^{-1} d_k d_m = 0.$$

So the new connection polynomial produces the sequence $\{S_1, S_2, \ldots, S_k\}$ with no discrepancy.

Here we present the algorithm:

Data: $S_1, S_2, \ldots, S_N$
$L = 0$ (the current length of LRS);
$\Lambda^{[k]}(x) = 1$ (the current connection polynomial);
$\Lambda^{[m-1]}(x) = 1$ (the connection polynomial before last length change);
$p = 1$ (p is $k-m$, the amount of shift in update);
$d_m = 1$ (previous discrepancy);
for k=*1 to* N **do**
  $d = S_k + \sum_{i=1}^{L} \Lambda_i^{[k]} S_{k-i}$ (compute discrepancy);
  if d=*0* **then**
    $p = p + 1$ (no change in polynomial);
  else
    if $2L \geq k$ **then**
      $\Lambda^{[k]}(x) = \Lambda^{[k]}(x) - d\, d_m^{-1} x^p \Lambda^{[m-1]}(x)$;
      $p = p + 1$;
    end
    $p(x) = \Lambda^{[k]}(x)$ (temporary storage);
    $\Lambda^{[k]}(x) = \Lambda^{[k]}(x) - d\, d_m^{-1} x^p \Lambda^{[m-1]}(x)$;
    $L = k - L$;
    $d_m = d$;
    $p = 1$; $\Lambda^{[m-1]}(x) = p(x)$;
  end
end
Result: Connection Polynomial Λ^{N-1}

Algorithm 3: Massey's Algorithm

5.1.7 Forney's Algorithm

Having found the error-locator polynomial and its roots, there is still one more step for the non-binary BCH and RS codes, which is to find the error values. Let us return to the equation **??**. Knowing the error locators, obtained from the roots of the locator polynomial, it is straightforward to set up and solve a set of linear equations:

$$\begin{bmatrix} X_1 & X_2 & X_3 & \cdots & X_\nu \\ X_1^2 & X_2^2 & X_3^2 & \cdots & X_\nu^2 \\ \vdots & & & & \\ X_1^{2t} & X_2^{2t} & X_3^{2t} & \cdots & X_\nu^{2t} \end{bmatrix} \begin{bmatrix} e_{i_1} \\ e_{i_2} \\ \vdots \\ e_{i_\nu} \end{bmatrix} = \begin{bmatrix} S_1 \\ S_2 \\ \vdots \\ S_{2t} \end{bmatrix} \tag{5.11}$$

Forney's algorithm provides a fast and efficient way to solve this system. Before present the formula some definitions are necessary. A **syndrome polynomial** is defined as

$$S(x) = S_1 + S_2 x + S_3 x^2 + \cdots + S_{2t} x^{2t-1} = \sum_{j=0}^{2t-1} S_{j+1} x^j. \tag{5.12}$$

An **error-evaluator polynomial** $\Omega(x)$ is defined by

$$\Omega(x) = S(x)\Lambda(x) \mod x^{2t}. \tag{5.13}$$

This equation is called **key equation**.

Let $f(x) = f_0 + f_1 x + f_2 x^2 + \cdots + f_t x^t$ be a polynomial with coefficients in some field $\mathbb{F}$. The **formal derivate**

$$f'(x) = f_1 + 2f_2' x + 3f_3' x + \cdots + t f_t' x^{t-1}. \tag{5.14}$$

If $f(x) \in \mathbb{F}[x]$, where $\mathbb{F}$ is a field of characteristic 2, then $f'(x)$ has no odd-powered terms.

Definition 5.3: In Forney's algorithm the error values for a RS code are computed by

$$e_{i_k} = -\frac{\Omega(X_k^{k-1})}{\Lambda'(X_k^{-1})}, \tag{5.15}$$

where $\Lambda'(x)$ is the formal derivative of $\Lambda(x)$.

5.2 A 'behavioral' decoder

Before start to describe the procedures of the section, consider the following important definition and properties.

Definition 5.4: aLet $Q(x,y)=\sum a_i x^{n_i} y^{m_i} \in \mathbb{F}[x,y]$. We define the degree of Q as $\deg Q = \max\{n_i + m_i\}$, $a_i \neq 0$.

Remark 5.5: Consider the polynomial $p(x) \in \mathbb{F}[x]$, where x is the operator σ^m. We have that:

$$p(\sigma)x^k = \sigma^m x^k = x^{m+k} = x^m x^k = p(x)x^k.$$

We will denote by $r = (y_1,\ldots,y_n)$ the received word and by $d_i = (d_0(x_i),\ldots,d_M(x_i))$ the transmitted symbol i, i.e., the evaluation of the polynomial $m_k(x)$ in which coefficients are the original k data symbols.

Let start presenting an important Sudan's theorem that is necessary to understand the next procedures.

Theorem 5.6: Let $Q(x,y) \in \mathbb{F}[x,y]$ be a bivariate polynomial of weighted degree l such that $Q(x_i,y_i) = 0$, $i = 1,\ldots,n$. Let $r = (y_1,\ldots,y_n)$ be a received word. Denote the corresponding transmitted message polynomial by $m(x)$. If r contains less than $n-l$ errors then $y - m(x)$ divides $Q(x,y)$.

The main idea of Sudan's in [12] decoding approach is to construct a polynomial $Q(x,y)$ such that $Q(x_i,y_i) = 0$, constructed from the MPUM polynomial matrix. It makes sense to minimize the weighted degree of this polynomial as this maximizes the number of errors that can be corrected that way, according to Theorem 5.6. In the decoding process, all factors of the form $y - \tilde{m}(x)$ are subsequently extracted to produce a list of candidate polynomials $\tilde{m}(x)$ of degree $< k$.

M. Kuijper in [12] presented one solution to apply the result of Sudan using the system theory. The main idea is the following: given a set of points (x_i,y_i), $i = 1,\ldots,n$, we associate n trajectories $w_i : \mathbb{Z} \to \mathbb{F}^{M+1}$ for an appropriate choice of M, and we write the polynomial $Q(x,y) = \sum_{j=0}^{M} d_j(x) y^j$. Then she used the idea introduced by Willems in[30] and construct the MPUM $\mathfrak{B}$ for this trajectories. Then, if $R(s)$ is the matrix representation of the null space of $\mathfrak{B}$, we select a row $d(x)$ of minimal weighted row degree of R and finally we define $Q(x,y) = \sum_{j=0}^{M} d_j(x) y^j$, where the $d_i(x)s$ are the entries of $d(x)$. By Theorem 5.6, $Q(x,y)$

constructed in this way is a bivariate polynomial of minimal weighted degree that interpolates the data points (x_i, y_i).

Corollary 5.7: Let $\mathfrak{B}$ be the MPUM of $w_1^0, \ldots, w_n^0$ defined in Theorem 3.19 and M the integer number defined in equation 5.20. Let $R(x) \in \mathbb{F}^{(M+1)\times(M+1)}[x]$ be a weighted row reduced representation of $\mathfrak{B}$ and let $d(x) = [d_0(x) \cdots d_M(x)]$ be a row of $R(x)$ of minimal weighted degree. Define $Q(x,y) = \sum_{j=0}^{M} d_j(x) y^j$. Then $Q(x,y)$ is a polynomial of minimal weighted degree with $Q(x_i, y_i) = 0$ for $i = 1, \ldots, n$.

So lets start to construct the trajectories. We have that $Q(x_i, y_i) = 0$, i.e,

$$\begin{bmatrix} d_0(x_i) & \cdots & d_M(x_i) \end{bmatrix} \begin{bmatrix} 1 \\ y_i \\ \vdots \\ y_i^M \end{bmatrix} = 0, i = 1, \ldots, n. \tag{5.16}$$

Now, considering the remark 5.5, we can apply the shift operator σ to $d(x)$ as follows,

$$\begin{bmatrix} d_0(\sigma) & \cdots & d_M(\sigma) \end{bmatrix} \begin{bmatrix} 1 \\ y_i \\ \vdots \\ y_i^M \end{bmatrix} x_i^k = \begin{bmatrix} d_0(x_i) & \cdots & d_M(x_i) \end{bmatrix} \begin{bmatrix} 1 \\ y_i \\ \vdots \\ y_i^M \end{bmatrix} x_i^k, i = 1, \ldots, n. \tag{5.17}$$

Now from equation 5.16 we can write

$$\begin{bmatrix} d_0(x_i) & \cdots & d_M(x_i) \end{bmatrix} \begin{bmatrix} 1 \\ y_i \\ \vdots \\ y_i^M \end{bmatrix} x_i^k = 0. \tag{5.18}$$

So, given the n data points (x_i, y_i), $i = 1, \ldots, n$, we associate n trajectories $w_i : \mathbb{Z} \to \mathbb{F}^{M+1}$ defined as following

$$w_i(k) = \begin{bmatrix} 1 \\ y_i \\ \vdots \\ y_i^M \end{bmatrix} x_i^k. \tag{5.19}$$

Then we need to find the MPUM $\mathfrak{B}$ for these trajectories. First we define an integer M such that

$$M = \max\left\{ j \in \mathbb{N} | j \leq \frac{n}{k-1} \right\}. \tag{5.20}$$

Then we apply the following MPUM algorithm:

Data: interpolation data (x_i, y_i), for $i = 1, \ldots, n$; parameter k and M
Result: $R_n(x)$-Matrix of minimal weighted row degree
$R_0(x) = I_{M+1}$;

$$L_0 = \begin{bmatrix} 0 \\ k-1 \\ 2(k-1) \\ \vdots \\ M(k-1) \end{bmatrix} \text{(weighted row degrees of } R_1(x))$$

for *i=1 to n* **do**

$$Y_i = \begin{bmatrix} 1 & y_i & \ldots & y_i^M \end{bmatrix}^T;$$

$\Delta_i = R_{i-1}(x_i) Y_i$ (define the $i-th$ error trajectory);
$V_i(x) = (x - x_i) e_{j*} e_{j*}^T + \sum_{j \neq j*} e_j (\Delta_i(j*) e_j^T - \Delta_i(j) e_{j*}^T)$;
(e is the canonic basis of $\mathbb{F}^{M+1}$ and $j*$ is the smallest integer for which $L_{i-1}(j*)$ is minimal among $\{L_{i-1}(j) | \Delta_i(j) \neq 0\}$.)
$R_i(x) = V_i(x) R_{i-1}(x)$ (update matrix $R_i(x)$);
$L_i = L_{i-1} + b_{j*}$ (update vector L_i);

end

Algorithm 4: MPUM Detailed Algorithm

Theorem 5.8: Let $R_n(x)$ be the $(M+1)\times(M+1)$ polynomial matrix that results from applying algorithm 4 to the interpolation data (x_i, y_i) for $i = 1, \ldots, n$. Let $[d_0(x) \ldots d_M(x)]$ be a row of $R_n(x)$ of lowest weighted row degree, say L. Then $Q(x, y) = d_0(x) + d_1(x) y + \cdots + d_M(x) y^M$ is an interpolation solution of minimal weighted row degree L.

5.3 List Decoding

Now we will apply the above results in a practical example.

Consider the example4.20, and suppose that we have transmitted the same sequence

$$m = (\alpha, 1), \tag{5.21}$$

where α is a primitive element over $\mathbb{F}_{2^2}$, and the corresponding codeword is

$$r = \{\alpha, \alpha^2, 0, 1\}. \tag{5.22}$$

Now suppose that a single error have occurred during the transmission, and the received codeword was the following:

$$r = \{\alpha, 1, 0, 1\}. \tag{5.23}$$

The dimension of $\mathbb{F}_{2^2}$ is $k = 2$ so, according to equation (5.20), $M = 3$. The transmitted data expressed in pairs is given by

$$\{(x, y)\} = ((0, \alpha), (1, 1), (\alpha, 0), (\alpha^2, 1)). \tag{5.24}$$

The corresponding trajectories are

$$w_1(k) = \begin{bmatrix} 1 \\ \alpha \\ \alpha^2 \\ 1 \\ \alpha \end{bmatrix} 0^k; w_2(k) = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} 1^k; w_3(k) = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \alpha^k; w_4(k) = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} (\alpha^2)^k;$$

Running Algorithm 4 we obtain:

- Initialization:

$$L_0 = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}; R_0(x) = I_5;$$

- Iteration 1:

$$\Delta_1 = R_0(0)Y_1 = Y_1 = \begin{bmatrix} 1 \\ \alpha \\ \alpha^2 \\ 1 \\ \alpha \end{bmatrix};\ j* = 1; V_1(x) = \begin{bmatrix} x-0 & 0 & 0 & 0 & 0 \\ -\alpha & 1 & 0 & 0 & 0 \\ -\alpha^2 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 \\ -\alpha & 0 & 0 & 0 & 1 \end{bmatrix};$$

$$R_1(x) = V_1(x)R_0(x) = \begin{bmatrix} x-0 & 0 & 0 & 0 & 0 \\ -\alpha & 1 & 0 & 0 & 0 \\ -\alpha^2 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 \\ -\alpha & 0 & 0 & 0 & 1 \end{bmatrix}; L_1 = L_0 + e_1 = \begin{bmatrix} 1 \\ 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$$

- Iteration 2:

$$\Delta_2 = R_1(1)Y_2 = \begin{bmatrix} 1 \\ \alpha^2 \\ \alpha \\ 0 \\ \alpha^2 \end{bmatrix};\ j* = 1; V_2(x) = \begin{bmatrix} x-1 & 0 & 0 & 0 & 0 \\ -\alpha^2 & 1 & 0 & 0 & 0 \\ -\alpha & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ -\alpha^2 & 0 & 0 & 0 & 1 \end{bmatrix};$$

$$R_2(x) = V_2(x)R_1(x) = \begin{bmatrix} x(x+1) & 0 & 0 & 0 & 0 \\ \alpha^2 x + \alpha & 1 & 0 & 0 & 0 \\ \alpha & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ \alpha^2 x + \alpha & 0 & 0 & 0 & 1 \end{bmatrix}; L_2 = L_1 + e_1 = \begin{bmatrix} 2 \\ 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$$

- Iteration 3:

$$\Delta_3 = R_2(\alpha)Y_3 = \begin{bmatrix} 1 \\ \alpha^2 \\ 0 \\ 1 \\ \alpha^2 \end{bmatrix}; j* = 2; V_3(x) = \begin{bmatrix} \alpha^2 & 1 & 0 & 0 & 0 \\ 0 & x+\alpha & 0 & 0 & 0 \\ 0 & 0 & \alpha^2 & 0 & 0 \\ 0 & 1 & 0 & \alpha^2 & 0 \\ 0 & \alpha^2 & 0 & 0 & \alpha^2 \end{bmatrix};$$

$$R_3(x) = V_3(x)R_2(x) = \begin{bmatrix} \alpha^2 x^2+\alpha & 1 & 0 & 0 & 0 \\ \alpha^2 x^2+x+\alpha^2 & x+\alpha & 0 & 0 & 0 \\ x+\alpha & 0 & \alpha^2 & 0 & 0 \\ \alpha^2 x+1 & 1 & 0 & \alpha^2 & 0 \\ 0 & \alpha^2 & 0 & 0 & \alpha^2 \end{bmatrix}; L_3 = L_2 + e_2 = \begin{bmatrix} 2 \\ 2 \\ 2 \\ 3 \\ 4 \end{bmatrix}$$

- Iteration 4:

$$\Delta_4 = R_3(\alpha)Y_4 = \begin{bmatrix} \alpha \\ \alpha^2 \\ \alpha \\ 1 \\ 0 \end{bmatrix}; j* = 1; V_4(x) = \begin{bmatrix} x+\alpha^2 & 0 & 0 & 0 & 0 \\ 1 & \alpha & 0 & 0 & 0 \\ \alpha & 0 & \alpha & 0 & 0 \\ 1 & 0 & 0 & \alpha & 0 \\ 0 & 0 & 0 & 0 & \alpha \end{bmatrix};$$

$$R_4(x) = V_4(x)R_3(x) = \begin{bmatrix} \alpha^2 x^3+\alpha x^2+\alpha x+1 & x+\alpha^2 & 0 & 0 & 0 \\ \alpha x^2+x+\alpha^2 & \alpha x+\alpha & 0 & 0 & 0 \\ x^2+\alpha x & \alpha & 1 & 0 & 0 \\ \alpha^2 x^2+x & \alpha^2 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

According to Definition 3.7, we need to construct a polynomial matrix $M(x,y) = R_n(x)N(y)$, such that:

$$N(y) = \text{diag}(1, y, y^2, y^3, y^4).$$

Thus we get:

$$M(x,y) = \begin{bmatrix} \alpha^2x^3+\alpha x^2+\alpha x+1 & x+\alpha^2 & 0 & 0 & 0 \\ \alpha x^2+x+\alpha^2 & \alpha x+\alpha & 0 & 0 & 0 \\ x^2+\alpha x & \alpha & 1 & 0 & 0 \\ \alpha^2x^2+x & \alpha^2 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & y & 0 & 0 & 0 \\ 0 & 0 & y^2 & 0 & 0 \\ 0 & 0 & 0 & y^3 & 0 \\ 0 & 0 & 0 & 0 & y^4 \end{bmatrix}$$

$$= \begin{bmatrix} \alpha^2x^3+\alpha x^2y+\alpha x+1 & (x+\alpha^2)y & 0 & 0 & 0 \\ \alpha x^2+x+\alpha^2 & (\alpha x+\alpha)y & 0 & 0 & 0 \\ x^2+\alpha x & \alpha y & y^2 & 0 & 0 \\ \alpha^2x^2+x & \alpha^2y & 0 & y^3 & 0 \\ 0 & y & 0 & 0 & y^4 \end{bmatrix}$$

Hence, according Definition 5.4, the row degrees of $M(x,y)$ are $4,2,2,3$ and 4 respectively. The second and third rows have the minimal weight degree. Now we construct the polynomials $M_(x,y) = \sum_{j=0}^{M} r_j(x)y^j$, where $r_i(x)$ are the entries of the row of minimal weighted degree of $R(x)$. According [12], it turns out that $M(x,y)$ constructed in this way is a bivariate polynomial of minimal weighted degree that interpolates the data point (x_i, y_i), $i = 1,\ldots,n$. It turns out that:

$$M_2(x,y) = \alpha x^2+x+\alpha^2+\alpha xy+\alpha y,$$

and $M_2(x,y) = 0$, so

$$\alpha x^2+x+\alpha^2+\alpha xy+\alpha y = 0. \tag{5.25}$$

Considering $y = x+b$, for some $b \in \mathbb{F}^4$, we have:

$$\begin{aligned} &\alpha x^2+x+\alpha^2+\alpha xy+\alpha y = 0 \\ &\Leftrightarrow \alpha x^2+x+\alpha^2+\alpha x(x+b)+\alpha(x+b) = 0 \\ &\Leftrightarrow \alpha x^2+x+\alpha^2+\alpha x^2+\alpha xb+\alpha x+\alpha b = 0 \\ &\Leftrightarrow b(\alpha x+\alpha) = \alpha(x\alpha+\alpha) \\ &\Leftrightarrow b = \alpha. \end{aligned}$$

Thus $y = x = m(x)$. Also we have that

$$M_3(x,y) = x^2+\alpha x+\alpha y+y^2 = 0.$$

Considering $y = x + b$, for some $b \in \mathbb{F}^4$, we have:

$$\begin{aligned} x^2 + \alpha x + \alpha y + y^2 = 0 &\Leftrightarrow x^2 + \alpha x + \alpha(x+b) + (x+b)^2 = 0 \\ &\Leftrightarrow x^2 + \alpha x + \alpha x + \alpha b + x^2 + 2xb + b^2 = 0 \\ &\Leftrightarrow b(\alpha + b) = 0 \\ &\Leftrightarrow b = 0 \text{ or } b = \alpha. \end{aligned}$$

We have now two possibilities, $y = x = m(x)$ or $y = x + \alpha = m(x)$.

The next steps are the following: first we convert the possible messages in a codeword form by evaluating them in all elements of the field as in 4.20. Then original codeword will be the one that has the minimum distance to the received word.

Let $m_1 = (0,1) \rightarrow m_1(x) = x$ and $m_2 = (\alpha, 1) \rightarrow m_2(x) = \alpha + x$. The possible code words will be:

$$c_1 = (m_1(0), m_1(1), m_1(\alpha), m_1(\alpha^2)) = (0, 1, \alpha, \alpha^2)$$

and

$$c_2 = (m_2(0), m_2(1), m_2(\alpha), m_2(\alpha^2)) = (\alpha, \alpha^2, 0, 1)$$

And therefore $d_{min}(c_1, r) = 3$ and $d_{min}(c_2, r) = 1$. It means that $m = (\alpha, 1)$ is the original word.

Chapter 6

Conclusions and Future Work

In this work we have presented a system theoretic approach to list decoding using the concept of behavior. The contribution lies in the behavior solution to the bivariate interpolation problem associated to the decoding problem. With the received word a set of trajectories is associated. These trajectories in turn generate a behavior. This behavior may be represented as the kernel of a matrix of polynomials in the shift. After transforming this matrix into weighted row reduced form a row of minimal weighted row degree is selected. Finally, the interpolation bivariate polynomial is obtained from that row. At each step of this procedure, weighted row reduceness is guaranteed so that the transformation to weighted row reduced at the end is needless. An algorithm for this is presented in[15]. Although we only need one row of minimal weighted row degree, we compute the complete weighted row reduced representation of the behavior. One question that stays on the air is what additional information about the transmitted codeword is possibly carried by the other rows. Moreover, it may be interesting explore the parallelism between the traditional decoding technics and this new approach, by the efficiency point of view. What are the advantages and disadvantages that we obtain if we choose this new method for decoding, is a good question for a future work.

Bibliography

[1] M. Ali and M. Kuijper. A parametric approach to list decoding of reed-solomon codes using interpolation. *IEEE Transactions on Information Theory*, 57(10):6718–6728, 2011.

[2] R. Blahut. *Algebraic Codes for Data Transmission*. Cambridge University Press, 2003.

[3] R. C. Bose and D. K. Ray-Chaudhuri. On a class of error correcting binary group codes. *Information and Control 3*, pages 68–79, 1960.

[4] N. J. A. S. F. J. MacWilliams and J. M. Goethals. The macwilliams identities for nonlinear codes, 1971.

[5] A. Hocquenghem. Codes correcteurs d'erreurs. *Chiffres*.

[6] M. Kuijper. An algorithm for constructing a minimal partial realization in the multivariable case. *Systems and Control Letters*, pages 225–233, 1997.

[7] M. Kuijper. Further results on the use of a generalized b-m algorithm for bch decoding beyond the designed error-correcting capability. *Proceecing of the 13th Symposium on Apllied Algebra Algebraic Algorithms, and Error-Correcting Codes (AAECC)*, pages 98–99, 1999.

[8] M. Kuijper. *Behavioral interpolation for coding and control*, volume 3. 2000.

[9] M. Kuijper. Algorithms for decoding and interpolation. *Brian Marcus and Joaquim Rosenthal*, 123:265–282, 2001.

[10] M. Kuijper, R. Pinto, and J. W. Polderman. The predictable degree property and row reducedness for systems over a finite ring. *Linear Algebra and its Applications*, 425(2–3):776 – 796, 2007.

[11] M. Kuijper and J. W. Polderman. A behavioral approach to list decoding. In *Proc. 15th Int. Symp. Mathematical Theory of Networks and Systems*, pages 1–13, 2002.

[12] M. Kuijper and J. W. Polderman. Behavioral models for list decoding. *Mathematical and Computer Modelling of Dynamical Systems*, 8(4):429–443, 2002.

[13] M. Kuijper and J. W. Polderman. Reed-solomon list decoding from a system theoretic perspective. *IEEE Transactions on Information Theory*, 50:259–271, 2004.

[14] M. Kuijper and J. W. Polderman. Systems theoretic methods in decoding. In R. N. J. Veldhuis, H. S. Cronie, and F. W. Hoeksema, editors, *28th Symposium on Information Theory in the Benelux*, pages 19–26, Enschede, May 2007. Werkgemeenschap voor Informatie- en Communicatietheorie.

[15] M. Kuijper and Willems. Behavioral models for list decoding. *Mathematical and Computer Modelling of Dynamical Systems*, 8(4):429–443, 2002.

[16] M. Kuijper and J. C. Willems. On constructing a shortest linear recurrence relation. *Automatic Control, IEEE Transactions on*, 42(11):1554–1558, 1997.

[17] R. Lidl and H. Niederreiter. *Finite fields*. Cambridge University Press, 1997.

[18] J. L. Massey and M. K. Sain. Codes, automata and continuous systems: Explicit interconnections. *IEEE Trans. on Auto. Cont.rol*, 23:33–38, 1967.

[19] T. K. Moon. *Error Correction Coding: Mathematical Methods and Algorithms*. Wiley-Interscience, 2005.

[20] J. W. Polderman and J. C. Willems. *Introduction to Mathematical Systems Theory: A Behavioral Approach*, volume 26 of *Texts in Applied Mathematics*. Springer-Verlag, New York, 1998.

[21] I. Reed. A class of multiple-error-correcting codes and the decoding scheme. *Information Theory, Transactions of the IRE Professional Group on*, 4(4):38–49, september 1954.

[22] I. Reed and S. Golomb. Polynomial codes over certain finite fields. *Joint Society of Industrial and Applied Mathematics Journal*, 8(2):300–304, June 1960.

[23] J. Rosenthal, J. M. Schumacher, and E. V. York. The Behavior Of Convolution Codes. Technical report, 1995.

[24] C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, (3), 1948.

[25] C. E. Shannon. A mathematical theory of communication. *Bell System Tech. J.*, 27:379–423, 623–656, 1948.

[26] C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423, 1948.

[27] M. Sudan. Decoding of reed solomon codes beyond the error-correction bound. *Journal of Complexity*, 13(1):180–193, 1997.

[28] A. S. Tanenbaum. *Computer Network*. Prentice Hall, 2002.

[29] J. H. Venkatesan Guruswami and S. Koppaety. On the list-decodability of random linear codes. 2010.

[30] J. C. Willems. From time series to linear system-part i. *Pergamon Journals Ltd*, 22(5):561–580, 1986.

[31] J. C. Willems. From time series to linear system-part ii. *Pergamon Journals Ltd*, 22(5), 1986.